

**Multiplicity dependence of inclusive  $p_t$  spectra from  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV**

J. Adams,<sup>2</sup> M. M. Aggarwal,<sup>29</sup> Z. Ahammed,<sup>44</sup> J. Amonett,<sup>19</sup> B. D. Anderson,<sup>19</sup> M. Anderson,<sup>6</sup> D. Arkhipkin,<sup>12</sup> G. S. Averichev,<sup>11</sup> Y. Bai,<sup>27</sup> J. Balewski,<sup>16</sup> O. Barannikova,<sup>2</sup> L. S. Barnby,<sup>2</sup> J. Baudot,<sup>17</sup> S. Bekele,<sup>28</sup> V. V. Belaga,<sup>11</sup> A. Bellingeri-Laurikainen,<sup>39</sup> R. Bellwied,<sup>47</sup> F. Benedosso,<sup>27</sup> B. I. Bezverkhny,<sup>49</sup> S. Bhardwaj,<sup>34</sup> A. Bhasin,<sup>18</sup> A. K. Bhati,<sup>29</sup> H. Bichsel,<sup>46</sup> J. Bielcik,<sup>49</sup> J. Bielcikova,<sup>49</sup> L. C. Bland,<sup>3</sup> C. O. Blyth,<sup>2</sup> S.-L. Blyth,<sup>21</sup> B. E. Bonner,<sup>35</sup> M. Botje,<sup>27</sup> J. Bouchet,<sup>39</sup> A. V. Brandin,<sup>25</sup> A. Bravar,<sup>3</sup> M. Bystersky,<sup>10</sup> R. V. Cadman,<sup>1</sup> X. Z. Cai,<sup>38</sup> H. Caines,<sup>49</sup> M. Calderón de la Barca Sánchez,<sup>6</sup> J. Castillo,<sup>27</sup> O. Catu,<sup>49</sup> D. Cebra,<sup>6</sup> Z. Chajecki,<sup>28</sup> P. Chaloupka,<sup>10</sup> S. Chattopadhyay,<sup>44</sup> H. F. Chen,<sup>37</sup> J. H. Chen,<sup>38</sup> Y. Chen,<sup>7</sup> J. Cheng,<sup>42</sup> M. Cherney,<sup>9</sup> A. Chikanian,<sup>49</sup> H. A. Choi,<sup>33</sup> W. Christie,<sup>3</sup> J. P. Coffin,<sup>17</sup> T. M. Cormier,<sup>47</sup> M. R. Cosentino,<sup>36</sup> J. G. Cramer,<sup>46</sup> H. J. Crawford,<sup>5</sup> D. Das,<sup>44</sup> S. Das,<sup>44</sup> M. Daugherty,<sup>41</sup> M. M. de Moura,<sup>36</sup> T. G. Dedovich,<sup>11</sup> M. DePhillips,<sup>3</sup> A. A. Derevschikov,<sup>31</sup> L. Didenko,<sup>3</sup> T. Dietel,<sup>13</sup> P. Djawotho,<sup>16</sup> S. M. Dogra,<sup>18</sup> W. J. Dong,<sup>7</sup> X. Dong,<sup>37</sup> J. E. Draper,<sup>6</sup> F. Du,<sup>49</sup> V. B. Dunin,<sup>11</sup> J. C. Dunlop,<sup>3</sup> M. R. Dutta Mazumdar,<sup>44</sup> V. Eckardt,<sup>23</sup> W. R. Edwards,<sup>21</sup> L. G. Efimov,<sup>11</sup> V. Emelianov,<sup>25</sup> J. Engelage,<sup>5</sup> G. Eppley,<sup>35</sup> B. Erasmus,<sup>39</sup> M. Estienne,<sup>17</sup> P. Fachini,<sup>3</sup> R. Fatemi,<sup>22</sup> J. Fedorisin,<sup>11</sup> K. Filimonov,<sup>21</sup> P. Filip,<sup>12</sup> E. Finch,<sup>49</sup> V. Fine,<sup>3</sup> Y. Fisyak,<sup>3</sup> J. Fu,<sup>48</sup> C. A. Gagliardi,<sup>40</sup> L. Gaillard,<sup>2</sup> J. Gans,<sup>49</sup> M. S. Ganti,<sup>44</sup> V. Ghazikhanian,<sup>7</sup> P. Ghosh,<sup>44</sup> J. E. Gonzalez,<sup>7</sup> Y. G. Gorbunov,<sup>9</sup> H. Gos,<sup>45</sup> O. Grebenyuk,<sup>27</sup> D. Grosnick,<sup>43</sup> S. M. Guertin,<sup>7</sup> K. S. F. F. Guimaraes,<sup>36</sup> Y. Guo,<sup>47</sup> N. Gupta,<sup>18</sup> T. D. Gutierrez,<sup>6</sup> B. Haag,<sup>6</sup> T. J. Hallman,<sup>3</sup> A. Hamed,<sup>47</sup> J. W. Harris,<sup>49</sup> W. He,<sup>16</sup> M. Heinz,<sup>49</sup> T. W. Henry,<sup>40</sup> S. Heppelmann,<sup>30</sup> B. Hippolyte,<sup>17</sup> A. Hirsch,<sup>32</sup> E. Hjort,<sup>21</sup> G. W. Hoffmann,<sup>41</sup> M. J. Horner,<sup>21</sup> H. Z. Huang,<sup>7</sup> S. L. Huang,<sup>37</sup> E. W. Hughes,<sup>4</sup> T. J. Humanic,<sup>28</sup> G. Igo,<sup>7</sup> P. Jacobs,<sup>21</sup> W. W. Jacobs,<sup>16</sup> P. Jakl,<sup>10</sup> F. Jia,<sup>20</sup> H. Jiang,<sup>7</sup> P. G. Jones,<sup>2</sup> E. G. Judd,<sup>5</sup> S. Kabana,<sup>39</sup> K. Kang,<sup>42</sup> J. Kapitan,<sup>10</sup> M. Kaplan,<sup>8</sup> D. Keane,<sup>19</sup> A. Kechechyan,<sup>11</sup> V. Yu. Khodyrev,<sup>31</sup> B. C. Kim,<sup>33</sup> J. Kiryluk,<sup>22</sup> A. Kisiel,<sup>45</sup> E. M. Kislov,<sup>11</sup> S. R. Klein,<sup>21</sup> D. D. Koetke,<sup>43</sup> T. Kollegger,<sup>13</sup> M. Kopytine,<sup>19</sup> L. Kotchenda,<sup>25</sup> V. Kouchpil,<sup>10</sup> K. L. Kowalik,<sup>21</sup> M. Kramer,<sup>26</sup> P. Kravtsov,<sup>25</sup> V. I. Kravtsov,<sup>31</sup> K. Krueger,<sup>1</sup> C. Kuhn,<sup>17</sup> A. I. Kulikov,<sup>11</sup> A. Kumar,<sup>29</sup> A. A. Kuznetsov,<sup>11</sup> M. A. C. Lamont,<sup>49</sup> J. M. Landgraf,<sup>3</sup> S. Lange,<sup>13</sup> S. LaPointe,<sup>47</sup> F. Laue,<sup>3</sup> J. Lauret,<sup>3</sup> A. Lebedev,<sup>3</sup> R. Lednicky,<sup>12</sup> C.-H. Lee,<sup>33</sup> S. Lehocka,<sup>11</sup> M. J. LeVine,<sup>3</sup> C. Li,<sup>37</sup> Q. Li,<sup>47</sup> Y. Li,<sup>42</sup> G. Lin,<sup>49</sup> S. J. Lindenbaum,<sup>26</sup> M. A. Lisa,<sup>28</sup> F. Liu,<sup>48</sup> H. Liu,<sup>37</sup> J. Liu,<sup>35</sup> L. Liu,<sup>48</sup> Z. Liu,<sup>48</sup> T. Ljubicic,<sup>3</sup> W. J. Llope,<sup>35</sup> H. Long,<sup>7</sup> R. S. Longacre,<sup>3</sup> M. Lopez-Noriega,<sup>28</sup> W. A. Love,<sup>3</sup> Y. Lu,<sup>48</sup> T. Ludlam,<sup>3</sup> D. Lynn,<sup>3</sup> G. L. Ma,<sup>38</sup> J. G. Ma,<sup>7</sup> Y. G. Ma,<sup>38</sup> D. Magestro,<sup>28</sup> D. P. Mahapatra,<sup>14</sup> R. Majka,<sup>49</sup> L. K. Mangotra,<sup>18</sup> R. Manweiler,<sup>43</sup> S. Margetis,<sup>19</sup> C. Markert,<sup>19</sup> L. Martin,<sup>39</sup> H. S. Matis,<sup>21</sup> Yu. A. Matulenko,<sup>31</sup> C. J. McClain,<sup>1</sup> T. S. McShane,<sup>9</sup> Yu. Melnick,<sup>31</sup> A. Meschanin,<sup>31</sup> M. L. Miller,<sup>22</sup> N. G. Minaev,<sup>31</sup> S. Mioduszewski,<sup>40</sup> C. Mironov,<sup>19</sup> A. Mischke,<sup>27</sup> D. K. Mishra,<sup>14</sup> J. Mitchell,<sup>35</sup> B. Mohanty,<sup>44</sup> L. Molnar,<sup>32</sup> C. F. Moore,<sup>41</sup> D. A. Morozov,<sup>31</sup> M. G. Munhoz,<sup>36</sup> B. K. Nandi,<sup>15</sup> C. Nattrass,<sup>49</sup> T. K. Nayak,<sup>44</sup> J. M. Nelson,<sup>2</sup> P. K. Netrakanti,<sup>44</sup> V. A. Nikitin,<sup>12</sup> L. V. Nogach,<sup>31</sup> S. B. Nurushev,<sup>31</sup> G. Odyniec,<sup>21</sup> A. Ogawa,<sup>3</sup> V. Okorokov,<sup>25</sup> M. Oldenburg,<sup>21</sup> D. Olson,<sup>21</sup> M. Pachr,<sup>10</sup> S. K. Pal,<sup>44</sup> Y. Panebratsev,<sup>11</sup> S. Y. Panitkin,<sup>3</sup> A. I. Pavlinov,<sup>47</sup> T. Pawlak,<sup>45</sup> T. Peitzmann,<sup>27</sup> V. Perevoztchikov,<sup>3</sup> C. Perkins,<sup>5</sup> W. Peryt,<sup>45</sup> V. A. Petrov,<sup>47</sup> S. C. Phatak,<sup>14</sup> R. Picha,<sup>6</sup> M. Planinic,<sup>50</sup> J. Pluta,<sup>45</sup> N. Poljak,<sup>50</sup> N. Porile,<sup>32</sup> J. Porter,<sup>46</sup> A. M. Poskanzer,<sup>21</sup> M. Potekhin,<sup>3</sup> E. Potrebenikova,<sup>11</sup> B. V. K. S. Potukuchi,<sup>18</sup> D. Prindle,<sup>46</sup> C. Pruneau,<sup>47</sup> J. Putschke,<sup>21</sup> G. Rakness,<sup>30</sup> R. Raniwala,<sup>34</sup> S. Raniwala,<sup>34</sup> R. L. Ray,<sup>41</sup> S. V. Razin,<sup>11</sup> J. Reinnarth,<sup>39</sup> D. Relyea,<sup>4</sup> F. Retiere,<sup>21</sup> A. Ridiger,<sup>25</sup> H. G. Ritter,<sup>21</sup> J. B. Roberts,<sup>35</sup> O. V. Rogachevskiy,<sup>11</sup> J. L. Romero,<sup>6</sup> A. Rose,<sup>21</sup> C. Roy,<sup>39</sup> L. Ruan,<sup>21</sup> M. J. Russcher,<sup>27</sup> R. Sahoo,<sup>14</sup> I. Sakrejda,<sup>21</sup> S. Salur,<sup>49</sup> J. Sandweiss,<sup>49</sup> M. Sarsour,<sup>40</sup> P. S. Sazhin,<sup>11</sup> J. Schambach,<sup>41</sup> R. P. Scharenberg,<sup>32</sup> N. Schmitz,<sup>23</sup> K. Schweda,<sup>21</sup> J. Seger,<sup>9</sup> I. Selyuzhenkov,<sup>47</sup> P. Seyboth,<sup>23</sup> A. Shabetai,<sup>21</sup> E. Shahaliev,<sup>11</sup> M. Shao,<sup>37</sup> M. Sharma,<sup>29</sup> W. Q. Shen,<sup>38</sup> S. S. Shimanskiy,<sup>11</sup> E. Sichtermann,<sup>21</sup> F. Simon,<sup>22</sup> R. N. Singaraju,<sup>44</sup> N. Smirnov,<sup>49</sup> R. Snellings,<sup>27</sup> G. Sood,<sup>43</sup> P. Sorensen,<sup>3</sup> J. Sowinski,<sup>16</sup> J. Speltz,<sup>17</sup> H. M. Spinka,<sup>1</sup> B. Srivastava,<sup>32</sup> A. Stadnik,<sup>11</sup> T. D. S. Stanislaus,<sup>43</sup> R. Stock,<sup>13</sup> A. Stolpovsky,<sup>47</sup> M. Strikhanov,<sup>25</sup> B. Stringfellow,<sup>32</sup> A. A. P. Suaide,<sup>36</sup> E. Sugarbaker,<sup>28</sup> M. Sumner,<sup>10</sup> Z. Sun,<sup>20</sup> B. Surrow,<sup>22</sup> M. Swanger,<sup>9</sup> T. J. M. Symons,<sup>21</sup> A. Szanto de Toledo,<sup>36</sup> A. Tai,<sup>7</sup> J. Takahashi,<sup>36</sup> A. H. Tang,<sup>3</sup> T. Tarnowsky,<sup>32</sup> D. Thein,<sup>7</sup> J. H. Thomas,<sup>21</sup> A. R. Timmins,<sup>2</sup> S. Timoshenko,<sup>25</sup> M. Tokarev,<sup>11</sup> T. A. Trainor,<sup>46</sup> S. Trentalange,<sup>7</sup> R. E. Tribble,<sup>40</sup> O. D. Tsai,<sup>7</sup> J. Ulery,<sup>32</sup> T. Ullrich,<sup>3</sup> D. G. Underwood,<sup>1</sup> G. Van Buren,<sup>3</sup> N. van der Kolk,<sup>27</sup> M. van Leeuwen,<sup>21</sup> A. M. Vander Molen,<sup>24</sup> R. Varma,<sup>15</sup> I. M. Vasilevski,<sup>12</sup> A. N. Vasiliev,<sup>31</sup> R. Vernet,<sup>17</sup> S. E. Vigdor,<sup>16</sup> Y. P. Viyogi,<sup>44</sup> S. Vokal,<sup>11</sup> S. A. Voloshin,<sup>47</sup> W. T. Waggoner,<sup>9</sup> F. Wang,<sup>32</sup> G. Wang,<sup>7</sup> J. S. Wang,<sup>20</sup> X. L. Wang,<sup>37</sup> Y. Wang,<sup>42</sup> J. W. Watson,<sup>19</sup> J. C. Webb,<sup>43</sup> G. D. Westfall,<sup>24</sup> A. Wetzler,<sup>21</sup> C. Whitten, Jr.,<sup>7</sup> H. Wieman,<sup>21</sup> S. W. Wissink,<sup>16</sup> R. Witt,<sup>49</sup> J. Wood,<sup>7</sup> J. Wu,<sup>37</sup> N. Xu,<sup>21</sup> Q. H. Xu,<sup>21</sup> Z. Xu,<sup>3</sup> P. Yepes,<sup>35</sup> I.-K. Yoo,<sup>33</sup> V. I. Yurevich,<sup>11</sup> W. Zhan,<sup>21</sup> H. Zhang,<sup>3</sup> W. M. Zhang,<sup>19</sup> Y. Zhang,<sup>37</sup> Z. P. Zhang,<sup>37</sup> Y. Zhao,<sup>37</sup> C. Zhong,<sup>38</sup> R. Zoulkarneev,<sup>12</sup> Y. Zoulkarneeva,<sup>12</sup> A. N. Zubarev,<sup>11</sup> and J. X. Zuo<sup>38</sup>

<sup>1</sup>Argonne National Laboratory, Argonne, Illinois 60439, USA

- <sup>2</sup>*University of Birmingham, Birmingham, United Kingdom*
- <sup>3</sup>*Brookhaven National Laboratory, Upton, New York 11973, USA*
- <sup>4</sup>*California Institute of Technology, Pasadena, California 91125, USA*
- <sup>5</sup>*University of California, Berkeley, California 94720, USA*
- <sup>6</sup>*University of California, Davis, California 95616, USA*
- <sup>7</sup>*University of California, Los Angeles, California 90095, USA*
- <sup>8</sup>*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
- <sup>9</sup>*Creighton University, Omaha, Nebraska 68178, USA*
- <sup>10</sup>*Nuclear Physics Institute AS CR, 250 68 Řež/Prague, Czech Republic*
- <sup>11</sup>*Laboratory for High Energy (JINR), Dubna, Russia*
- <sup>12</sup>*Particle Physics Laboratory (JINR), Dubna, Russia*
- <sup>13</sup>*University of Frankfurt, Frankfurt, Germany*
- <sup>14</sup>*Institute of Physics, Bhubaneswar 751005, India*
- <sup>15</sup>*Indian Institute of Technology, Mumbai, India*
- <sup>16</sup>*Indiana University, Bloomington, Indiana 47408, USA*
- <sup>17</sup>*Institut de Recherches Subatomiques, Strasbourg, France*
- <sup>18</sup>*University of Jammu, Jammu 180001, India*
- <sup>19</sup>*Kent State University, Kent, Ohio 44242, USA*
- <sup>20</sup>*Institute of Modern Physics, Lanzhou, China*
- <sup>21</sup>*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
- <sup>22</sup>*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA*
- <sup>23</sup>*Max-Planck-Institut für Physik, Munich, Germany*
- <sup>24</sup>*Michigan State University, East Lansing, Michigan 48824, USA*
- <sup>25</sup>*Moscow Engineering Physics Institute, Moscow, Russia*
- <sup>26</sup>*City College of New York, New York City, New York 10031, USA*
- <sup>27</sup>*NIKHEF and Utrecht University, Amsterdam, The Netherlands*
- <sup>28</sup>*Ohio State University, Columbus, Ohio 43210, USA*
- <sup>29</sup>*Panjab University, Chandigarh 160014, India*
- <sup>30</sup>*Pennsylvania State University, University Park, Pennsylvania 16802, USA*
- <sup>31</sup>*Institute of High Energy Physics, Protvino, Russia*
- <sup>32</sup>*Purdue University, West Lafayette, Indiana 47907, USA*
- <sup>33</sup>*Pusan National University, Pusan, Republic of Korea*
- <sup>34</sup>*University of Rajasthan, Jaipur 302004, India*
- <sup>35</sup>*Rice University, Houston, Texas 77251, USA*
- <sup>36</sup>*Universidade de Sao Paulo, Sao Paulo, Brazil*
- <sup>37</sup>*University of Science & Technology of China, Hefei 230026, China*
- <sup>38</sup>*Shanghai Institute of Applied Physics, Shanghai 201800, China*
- <sup>39</sup>*SUBATECH, Nantes, France*
- <sup>40</sup>*Texas A&M University, College Station, Texas 77843, USA*
- <sup>41</sup>*University of Texas, Austin, Texas 78712, USA*
- <sup>42</sup>*Tsinghua University, Beijing 100084, China*
- <sup>43</sup>*Valparaiso University, Valparaiso, Indiana 46383, USA*
- <sup>44</sup>*Variable Energy Cyclotron Centre, Kolkata 700064, India*
- <sup>45</sup>*Warsaw University of Technology, Warsaw, Poland*
- <sup>46</sup>*University of Washington, Seattle, Washington 98195, USA*
- <sup>47</sup>*Wayne State University, Detroit, Michigan 48201, USA*
- <sup>48</sup>*Institute of Particle Physics, CCNU (HZNU), Wuhan 430079, China*
- <sup>49</sup>*Yale University, New Haven, Connecticut 06520, USA*
- <sup>50</sup>*University of Zagreb, Zagreb, HR-10002, Croatia*

(Received 26 June 2006; published 16 August 2006)

We report measurements of transverse momentum  $p_t$  spectra for ten event multiplicity classes of  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV. By analyzing the multiplicity dependence we find that the spectrum shape can be decomposed into a part with amplitude proportional to multiplicity and described by a Lévy distribution on transverse mass  $m_t$ , and a part with amplitude proportional to multiplicity squared and described by a Gaussian distribution on transverse rapidity  $y_t$ . The functional forms of the two parts are nearly independent of event multiplicity. The two parts can be identified with the soft and hard components of a two-component model of  $p$ - $p$  collisions. This analysis then provides the first isolation of the hard component of the  $p_t$  spectrum as a distribution of simple form on  $y_t$ .

## I. INTRODUCTION

The structure of the inclusive  $p_t$  spectrum from relativistic nuclear collisions is affected by several aspects of collision dynamics and by the final-state hadronization process. Comparisons of  $p$ - $p$ ,  $d$ -Au, and Au-Au  $p_t$  spectra at Relativistic Heavy Ion Collider (RHIC) suggest that a form of color-deconfined matter has been created in Au-Au collisions [1,2]. Particle-production mechanisms which could determine spectrum structure include soft parton scattering followed by longitudinal or “string” fragmentation [3] and hard parton scattering followed by transverse fragmentation [4]. Other mechanisms could be significant. The structure of the  $p_t$  spectrum at some achievable level of precision may therefore be complex. A summary of efforts to unfold and interpret the structure of inclusive  $p_t$  spectra from ISR to Fermilab and SPS energies in the context of jet phenomenology and quantum chromodynamic (QCD) theory is provided in [5].

At RHIC energies hard parton scattering is expected to dominate the spectrum at larger  $p_t$  and to be significantly modified in  $A$ - $A$  collisions (jet quenching) [4,6]. But how does hard scattering contribute at smaller  $p_t$ ? How does it interact with thermal or “soft” particle production? Is there an “intermediate”  $p_t$  region [7] with its own unique production mechanisms? Those issues remain unresolved after much theoretical speculation and experimental measurement and provide a context for the present analysis applied to high-statistics  $p_t$  spectra from ten multiplicity classes of  $p$ - $p$  collisions. The multiplicity dependence offers new access to underlying particle-production mechanisms.

$p_t$  spectra from relativistic nuclear collisions are conventionally modeled by the *power-law* function [8], a form suggested by measured jet systematics and perturbative QCD (pQCD) expectations. At larger  $p_t$  the spectrum is expected to tend asymptotically to the power-law form  $p_t^{-n}$  [9]. The strict power-law form is then generalized to the function  $A/(1 + p_t/p_0)^n$ , having the expected pQCD dependence at larger  $p_t$  but transitioning to an approximate Maxwell-Boltzmann form at smaller  $p_t$ , consistent with expectations for thermal particle production. Although the power-law function has been previously applied to  $p$ - $p$  data with apparently good fit quality ( $\chi^2$  within expected limits) it has not been tested with the precision of recently acquired RHIC  $p$ - $p$  data. One can question the validity of its underlying assumptions. For instance, why should a single model function adequately describe spectra which may represent a mixture of several particle-production mechanisms?

Alternatively, a model function can be formulated in terms of the *two-component* model of nuclear collisions [10], which identifies “soft”  $p$ - $p$  collisions with no hard parton scatter and “semihard” collisions with at least one

significant parton scatter (i.e., producing distinguishable hadron fragments). According to the two-component model the minimum-bias distribution on event multiplicity  $n_{\text{ch}}$  can be decomposed into separate negative-binomial distributions (NBD) identified with soft and semihard event types. We then expect the *fraction* of events with a hard parton collision to increase monotonically with selected event multiplicity  $n_{\text{ch}}$ . Variation of  $p_t$  spectra with  $n_{\text{ch}}$  could then provide a basis for isolating soft and hard (and possibly other) components of inclusive spectra on a statistical basis, where the hard spectrum component refers to the fragment  $p_t$  spectrum for hard-scattered partons, and the soft component is the  $p_t$  spectrum for soft particle production.

In this analysis we first test the ability of the conventional power-law model function to represent the data. We then reconsider the data with no *a priori* assumptions. We attempt to describe all spectrum structure with the simplest algebraic model required by the data (e.g., “simple” in terms of parameter number and functional forms—cf. Eq. (4) and Sec. XI) and then to associate the model elements with possible particle-production mechanisms. We adopt two new analysis techniques: (1) We introduce transverse rapidity  $y_t$  [11,12] as an alternative to  $p_t$ .  $y_t$  has the advantage that spectrum structure associated with hard parton scattering and fragmentation is *more uniformly represented* on a logarithmic variable:  $y_t$  corresponds to variable  $\xi_p = \ln(p_{\text{parton}}/p_{\text{fragment}})$  conventionally used to describe parton fragmentation functions in elementary collisions [13]. A simple description of soft particle production is not compromised by the choice of transverse rapidity. (2) We introduce the running integral of the  $y_t$  spectrum, which substantially reduces statistical fluctuations relative to significant structure and therefore improves the precision of the analysis.

In this paper we present high-statistics  $p_t$  spectra for ten multiplicity classes from  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV. We use the conventional power-law model function to fit those spectra and assess the quality of that description. We then construct running integrals of the spectra on  $y_t$  and define a reference function common to all  $n_{\text{ch}}$  values and based on the Lévy distribution. We use that reference to extract difference spectra which contain the  $n_{\text{ch}}$ -dependent parts of the spectra in a more differential form. We find that the difference spectra have a simple structure: the major component is well described by a Gaussian distribution with fixed shape and with amplitude (relative to the reference) linearly proportional to the particle multiplicity. To simplify presentation we initially describe approximate relationships and optimized parameters without errors. We then return to a comprehensive discussion of the parameter system and its errors and consistency in Sec. VI. This

analysis is based on  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV observed with the STAR detector at the RHIC.

## II. $p_t$ AND $y_t$ SPECTRA

Data for this analysis in the form of inclusive  $p_t$  spectra for unidentified charged particles were obtained from nonsingle-diffractive (NSD)  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV triggered by a coincidence of two beam-beam counters (BBC) in  $3.3 < |\eta| < 5$  [1]. Charged particles were measured with the STAR Time Projection Chamber (TPC) and Central Trigger Barrel (CTB) [14]. Particle momenta were determined with a 0.5 T magnetic field parallel to the beam ( $z$ ) axis. Primary charged particles were represented by TPC tracks falling within the acceptance for this analysis— $2\pi$  azimuth, pseudorapidity  $|\eta| < 0.5$ , and  $0.2 < p_t < 6$  GeV/c—and satisfying track cuts described in [1]. The *observed* particle multiplicity in the acceptance is denoted by  $\hat{n}_{\text{ch}}$ , whereas the corrected and  $p_t$ -extrapolated *true* event multiplicity is denoted by  $n_{\text{ch}}$ . From  $3 \times 10^6$  NSD events individual  $p_t$  distributions were formed for 10 primary-particle multiplicity classes indexed by the observed multiplicity:  $\hat{n}_{\text{ch}} \in [1, \dots, 8, 9 + 10, 11 + 12]$ .

To eliminate backgrounds from event pileup each TPC primary-track candidate was independently required to match a CTB/trigger timing requirement (100 ns, matching efficiency 94%, false-coincidence background 2%) and project to the beam line within 1 cm transverse distance of closest approach. No other vertex requirement was applied to the primary tracks. The event-vertex  $z$  position was estimated by the arithmetic mean  $\bar{z}$  of projected track  $z$  for all CTB-matched primary tracks in an event. Events with  $|\bar{z}| < 75$  cm were accepted for further analysis. The event vertex was not included in primary-track  $p_t$  fits. That procedure eliminated pileup-event tracks, selected those events well positioned relative to the TPC, and minimized correlations of individual track  $p_t$  and  $p_t$  spectrum shape with event multiplicity or event triggering not related to collision dynamics.

The resulting  $p_t$  spectra were corrected for tracking efficiency, backgrounds, and momentum resolution. Tracking acceptance and efficiency on  $(p_t, \eta, z)$  and backgrounds were determined by embedding Hijing events in data events with at least one empty bunch (so-called abort-gap events). The same fractional correction was applied to all multiplicity classes. The correction factor was 1.45 at 0.2 GeV/c, falling to 1.2 at 0.5 GeV/c, and thereafter smoothly to 1 at 6 GeV/c. Efficiency- and acceptance-corrected (but not  $p_t$ -extrapolated) spectra integrate to multiplicity  $n'_{\text{ch}} = (1.35 \pm 0.015)\hat{n}_{\text{ch}}$ , while the corrected and  $p_t$ -extrapolated per-event spectra integrate to “true” multiplicity  $n_{\text{ch}} = (2.0 \pm 0.02)\hat{n}_{\text{ch}}$ . The errors reflect the spectrum-to-spectrum *relative* normalization uncertainties most relevant to this differential analysis. The normalization uncertainty common to all spectra is about 10%.

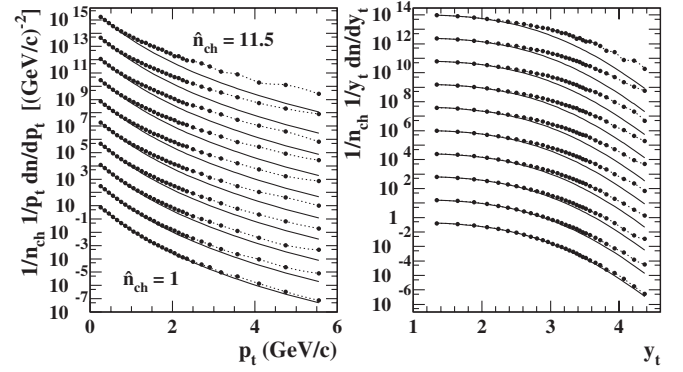


FIG. 1. Corrected and normalized charged-particle spectra on transverse momentum  $p_t$  (left) and transverse rapidity  $y_t$  (right) for 10 event multiplicity classes, displaced upward by successive factors 40 relative to  $\hat{n}_{\text{ch}} = 1$  at the bottom. Solid curves represent reference function  $n_s/n_{\text{ch}} \cdot S_0(y_t)$  (cf. Sec. IV C). Dotted curves are spline fits to guide the eye.

In Fig. 1 (left panel) corrected and normalized per-event  $p_t$  spectra are plotted as points in the form  $1/n_{\text{ch}} 1/p_t dn/dp_t$  for ten multiplicity classes, offset by successive factors 40 (except for  $\hat{n}_{\text{ch}} = 1$  at bottom). Parentheses for ratio prefactors of spectrum densities in the form  $dn/dx$  are omitted to lighten notation. In other cases ratio prefactors are separated from densities by a dot. Corrected and extrapolated spectra normalized by  $n_{\text{ch}}$  all integrate to unity in the sense of Eq. (1) for  $p_t$  or  $y_t$ , with the integration limit  $\rightarrow \infty$ . In Fig. 1 (right panel) equivalent spectra on transverse rapidity are plotted. Hard parton scattering leading to transverse fragmentation may be better described on transverse rapidity  $y_t = \ln\{(m_t + p_t)/m_0\}$ , with transverse mass  $m_t \equiv \sqrt{p_t^2 + m_0^2}$ , and pion mass  $m_\pi$  assumed for  $m_0$ :  $y_t = 2 \Rightarrow p_t \sim 0.5$  GeV/c and  $y_t = 4.5 \Rightarrow p_t \sim 6$  GeV/c. The solid curves  $n_s/n_{\text{ch}} \cdot S_0$  provide a visual reference for the data.  $n_s(\hat{n}_{\text{ch}})$  and  $S_0(p_t \text{ or } y_t)$  are defined below, and function  $S_0$  is by definition independent of  $\hat{n}_{\text{ch}}$ .

## III. POWER-LAW ANALYSIS

The *power-law* function is the conventional model function applied to  $p_t$  spectra from relativistic nuclear collisions [8]. Said to be “QCD-inspired,” the function  $A/(1 + p_t/p_0)^n$  goes asymptotically to  $p_t^{-n}$  at large  $p_t$  (hence “power-law”) and approximates an exponential at small  $p_t$ . The argument supporting the power-law function assumes that  $p_t$  spectra at larger collision energies can be modeled with a single functional form. In this part of the analysis we test that assumption. The  $p_t$  spectra for ten multiplicity classes in Fig. 1 were fitted with the three-parameter power-law model function defined above. Parameters  $A$ ,  $p_0$ , and  $n$  were independently varied to minimize  $\chi^2$  for each multiplicity class (in all fitting  $\chi^2$  was calculated using only statistical errors). The inclusive



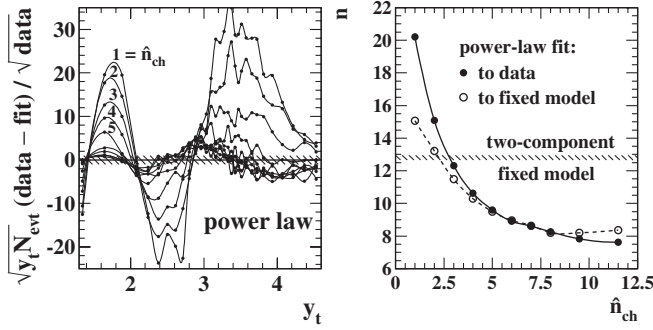


FIG. 2. Left: Relative residuals from power-law fits to  $p_t$  spectra in Fig. 1. The hatched band represents the expected statistical errors for STAR data. Right: Exponents  $n$  from power-law fits to data (solid points) and to corresponding two-component fixed-model functions (open circles, see Sec. VI) compared to the two-component fixed-model Lévy exponent  $12.8 \pm 0.15$  (hatched band).

mean  $p_t$  was extracted for each class as  $\langle p_t \rangle \equiv 2p_0/(n - 3)$  (cf. Sec. VII for those results).

In Fig. 2 (left panel) we plot *relative* fit residuals  $\sqrt{y_t N_{\text{evt}}}(\text{data} - \text{fit})/\sqrt{\text{data}}$  distributed on  $y_t$ . The points indicate the actual data positions. The quantity plotted insures that the residuals are directly measured in units of the root-mean-square statistical error at each  $y_t$ . These relative residuals are then similar to Pearson's correlation coefficient or relative covariance [15]. Poisson errors apply to  $dn/dy_t$ , whereas the spectra plotted in Fig. 1 ("data") are of the form  $1/y_t dn/dy_t$ . Thus, a factor  $\sqrt{y_t}$  is required to make the statistical reference uniform on  $y_t$  in these residuals plots. The residuals structure on  $p_t$  is equivalent to that on  $y_t$  within a Jacobian factor (the fits were actually done on  $p_t$  and the residuals transformed to  $y_t$  for this plot). As noted in the discussion in Sec. XI and elsewhere, much of the structure due to hard scattering and fragmentation is displaced to small  $p_t$  in a nonlinear way when plotted on  $p_t$ .

The large-wavelength residuals in Fig. 2 (left panel) exceed the expected statistical error (hatched band) by up to  $30\times$  and are similar in form for various  $\hat{n}_{\text{ch}}$  classes, revealing a large systematic disagreement between the power-law model and data. The small-wavelength structure, mainly attributable to true statistical fluctuations, is consistent with expectations (hatched band). The argument supporting the power-law model of  $p_t$  spectra is thus shown to fail when tested with high-statistics STAR  $p$ - $p$  data.

In Fig. 2 (right panel) we plot best-fit values of power-law exponent  $n$  vs  $\hat{n}_{\text{ch}}$  resulting from fits to data (solid points) and to the two-component model functions described later in this paper (open circles). The latter points and hatched band are discussed in Sec. XI. We observe a very strong variation of  $n$  with multiplicity. Reduction of  $n$  with increasing hard scattering is expected in the power-

law context, but we find that the physical mechanism is different from the theoretical expectation (cf. Sec. XI).

We observe very strong disagreement between the power-law model function and data, whereas a previous UA1 (SPPS) analysis reported power-law fits with reasonable  $\chi^2$  at the same energy [8]. The UA1 results are nevertheless consistent with the present analysis because that analysis was inclusive on  $n_{\text{ch}}$  and employed only 20 k minimum-bias events (vs  $3 \times 10^6$  for the present analysis). That analysis was therefore statistically insensitive to the structures apparent in Fig. 2. Statistics for the UA1 minimum-bias  $p_t$  spectrum are comparable to the  $\hat{n}_{\text{ch}} = 11.5$  multiplicity class in this study, but the latter contains about  $10\times$  the hard component in the UA1 minimum-bias spectrum. An E735 (FNAL) analysis of spectrometer data at 0.3, 0.55, 1.0, and 1.8 TeV [16], including multiplicity dependence of spectrum shapes, also obtained satisfactory power-law fits to  $p_t$  spectra. However the effective event number was comparable to the UA1 study, in part because of the reduced angular acceptance of the spectrometer relative to the STAR CTB detector, and the  $p_t$  acceptance [0.15, 3] GeV/c was considerably less than STAR or UA1, further reducing sensitivity to spectrum shape. Given this exclusion of the power-law model we now seek an alternative model which best describes  $p_t$  spectra from relativistic nuclear collisions.

#### IV. RUNNING INTEGRATION

Running integration provides substantial noise reduction for spectrum analysis, thereby improving precision. In this section we examine the  $n_{\text{ch}}$  dependence of differential and integrated spectra and define alternative normalization factor  $n_s(\hat{n}_{\text{ch}})$  and reference function  $S_0$ .

##### A. Spectrum normalization

In Fig. 3 (left panel) the spectra from Fig. 1 (right panel) are replotted without vertical offsets as spline curves for detailed comparison. No assumptions have been made about the data, and all spectra integrate to unity when extrapolated. The dash-dot curve is reference  $S_0$  defined in this section. To facilitate the discussion we identify three regions on  $y_t$  separated by the vertical dotted lines:  $A = [1.3, 1.9]$ ,  $B = [1.9, 3.4]$ , and  $C = [3.4, 4.5]$ . The region below  $y_t = 1.3$  is outside the  $p_t$  acceptance. Regions A and C are defined such that the curves within them are nearly constant relative to one another, whereas in region B the differences between curves vary rapidly.

The trend of the spectra with increasing  $\hat{n}_{\text{ch}}$  is counterbalancing changes within A and C: linear decrease in A (see inset) and linear increase in C. The relative variation in the two regions over the observed  $\hat{n}_{\text{ch}}$  range is quite different: 10% reduction in A and  $10\times$  increase in C. Such balancing variations are expected if the yield in C increases relative to A with  $\hat{n}_{\text{ch}}$ , due to the requirement that the normalized spectra must integrate to unity. We conclude

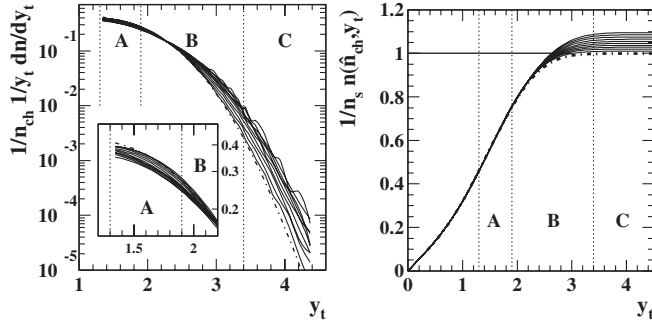


FIG. 3. Left: Spectra from Fig. 1 (right panel) replotted as spline curves and without offsets (solid curves) compared to reference  $S_0$  (dash-dot curve). Right: Running integrals Eq. (1) of extrapolated  $y_t$  spectra in Fig. 1 divided by  $n_s/n_{ch}$  (solid curves) compared to running integral  $N_0(y_t)$  of reference  $S_0(y_t)$  (dash-dot curve). The ten data curves from bottom to top correspond to increasing  $\hat{n}_{ch} \in [1, 11.5]$ .

that with increasing  $\hat{n}_{ch}$  additional particle yield localized on  $y_t$  and dominating region C is added to the spectrum.

The apparent reduction at smaller  $y_t$  is then a trivial effect of the unit-integral condition which can be compensated by changing the normalization. We normalize the spectra not by true total multiplicity  $n_{ch}$  but by multiplicity  $n_s$  defined such that the normalized spectra approximately coincide within region A. The variation of lower end point positions with  $\hat{n}_{ch}$  is compensated within errors by normalizing with the linear function  $\tilde{n}_s(\hat{n}_{ch}) = 2\hat{n}_{ch}(1 - 0.013\hat{n}_{ch})$  (function  $\tilde{n}_s$  estimates multiplicity  $n_s$ ). The negative term compensates the relative yield increase at larger  $y_t$ . The revised normalization also facilitates the running integration study described below.

### B. Running integrals and reference $S_0$

To calculate running integrals the measured spectra are extrapolated in the  $p_t$  interval  $[0, 0.2]$  GeV/c ( $y_t \in [0, 1.15]$ ) with reference function  $n_s/n_{ch} \cdot S_0$ . The extrapolation is relatively insensitive to the  $S_0$  parameters, insuring quick convergence of the  $S_0$  optimization procedure described below. The running integral of a  $y_t$  spectrum is defined by

$$n(\hat{n}_{ch}, y_t) = \int_0^{y_t} dy'_t y'_t \{1/y'_t dn(\hat{n}_{ch}, y'_t)/dy'_t\}. \quad (1)$$

In Fig. 3 (right panel) the normalized running integrals  $1/n_s(\hat{n}_{ch}) \cdot n(\hat{n}_{ch}, y_t)$  reveal the detailed structure of the spectra with much-improved signal-to-noise ratio. We observe that the integrals in the right panel indeed nearly coincide up to  $y_t \sim 2$ . Above that point (region B) the integrals separate. In region C the integrals all saturate, with nearly equal spacings between curves. That result provides a first detailed look at the localized (on  $y_t$ ) additional yield which produces the  $n_{ch}$  dependence of the  $y_t$  spectrum shape.

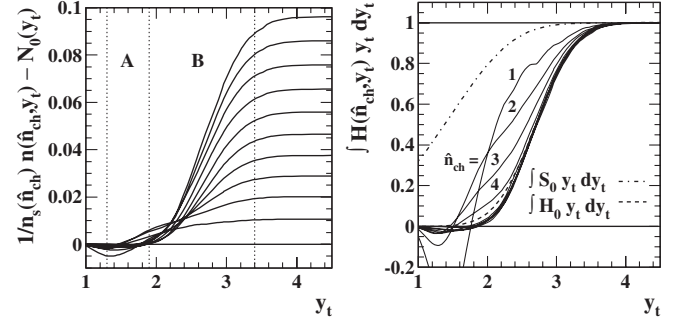


FIG. 4. Left: Differences between integrals of extrapolated  $y_t$  distributions in Fig. 1 according to Eq. (1), and integral  $N_0(y_t)$  of soft reference  $S_0(y_t)$ . The ten curves correspond to  $n_{ch} \in [1, 11.5]$  from bottom to top. Right: Distributions in the left panel divided by their end point values at  $y_t = 4.5$ . The dashed curve is the running integral of  $H_0$  (cf. Sec. V). The dash-dot curve is  $N_0$ , the running integral of  $S_0$ .

Given the results in Fig. 3 (right panel) the natural choice for a reference is one which coincides with all data curves for  $y_t < 2$  and defines a limiting case for the sequence of separated data curves at larger  $y_t$ . We therefore *define* the reference as the asymptotic limit of the  $y_t$  spectra (or their integrals) as  $\hat{n}_{ch} \rightarrow 0$ . For reasons discussed below we chose as a trial reference the Lévy distribution [17]

$$S_0(m_t; \beta_0, n) = A_s / (1 + \beta_0(m_t - m_0)/n)^n \quad (2)$$

defined on transverse mass  $m_t$  and suitably transformed to  $y_t$ .  $\beta_0 \equiv 1/T_0$  is an inverse-slope parameter. We find that the Lévy distribution with optimized parameters (dash-dot reference curves in Fig. 3) coincides with the desired asymptotic form. Determination of parameters  $n$  and  $\beta_0$  from the data is discussed in the next subsection. Amplitude  $A_s(\beta_0, n)$  is defined by the unit-integral normalization requirement on  $S_0$ .

The running integral of  $S_0$ , the dash-dot curve in Fig. 3 (right panel) denoted by  $N_0$ , is obtained by replacing the curly bracket in Eq. (1) with  $S_0(y_t)$ , in which case  $n(\hat{n}_{ch}, y_t) \rightarrow N_0(y_t)$  (also, see the legend in Fig. 4—right panel).  $N_0$  is thereby defined as the limit as  $\hat{n}_{ch} \rightarrow 0$  of the running integrals for the ten multiplicity classes. We can obtain a more differential picture by optimizing reference curve  $S_0$  and subtracting it and its running integral  $N_0$  from the data. Figure 4 (left panel) discussed in the next subsection reveals the  $n_{ch}$ -dependent yield increase as a localized structure on  $y_t$  and is used to optimize  $S_0$ . This differential procedure represents a new level of precision in spectrum analysis facilitated by the high-statistics STAR  $p$ - $p$  data and the running-integral technique.

### C. Optimizing reference $S_0$

In Fig. 4 (left panel) we plot the difference between running integrals  $1/n_s(\hat{n}_{ch}) \cdot n(\hat{n}_{ch}, y_t)$  of the corrected spectra in Fig. 3 (right panel) and reference integral

$N_0(y_t)$  (the dash-dot curve in that panel). In region  $B$  we observe a strong localized  $\hat{n}_{ch}$  dependence in the  $y_t$  spectra. The optimum parameters for  $S_0$  are derived as follows. Inverse-slope parameter  $\beta_0$  is adjusted to minimize residuals in region  $A$  of Fig. 4 (left panel).  $\beta_0$  determines the *average slope* of the residuals in that region. Exponent  $n$  then determines the size of the *first step* in region  $C$ .  $n$  is adjusted so that the first step follows the nearly linear trend of  $n_{ch}$  dependence in that  $y_t$  interval. Amplitude  $A_s(\beta, n)$  is determined by the unit-normalization requirement for  $S_0$ .

Changing either  $\beta_0$  or  $n$  in  $S_0$  *does not alter* the stepwise variation with  $\hat{n}_{ch}$  of the data curves in the left panel above the first step. That structure is inherent in the data and unaffected by the reference choice (cf. Fig. 3—right panel, before reference subtraction). The amplitude variation within region  $C$  is well represented by  $n_h/n_s = \alpha \hat{n}_{ch}$  with  $\alpha \sim 0.01$ , where  $n_h$  is the coefficient of  $H_0$  defined in the next subsection. That procedure determines reference  $S_0$  parameters  $A_s = 20.3 \pm 0.1$ ,  $n = 12.8 \pm 0.15$ , and  $T_0 = 0.1445 \pm 0.001$  GeV [18].

In Fig. 4 (right panel) the curves are obtained by dividing the curves in the left panel by their values at upper end point  $y_t = 4.5$  with approximate ratio  $n_h/n_s$ . Reference  $N_0(y_t)$  is included in the right panel as the dash-dot curve. Comparing  $N_0$  to the data integrals it is clear that the multiplicity dependence in Fig. 4 cannot be accommodated by adjusting  $S_0$ . With the exception of the first few  $\hat{n}_{ch}$  values (labeled curves) the integrals closely follow a common trend: an error function or running integral of a Gaussian which estimates *in a model-independent way* the running integral of the  $n_{ch}$ -independent model function  $H_0(y_t)$  determined differentially in the next section.

## V. DIFFERENTIAL ANALYSIS

Using running integrals we have defined a precision reference for the  $y_t$  spectra and isolated the  $n_{ch}$  dependence of those spectra relative to the reference. We now return to the differential  $y_t$  spectra and identify an additional spectrum component by subtracting the reference from the data. The dashed curve in Fig. 4 (right panel) (just visible near  $y_t = 2$ ) represents the running integral of model function  $H_0$  determined in this section.  $H_0(y_t)$  models the additional yield at larger  $y_t$  as a differential  $y_t$  spectrum component. It is already clear from Fig. 4 that the shape of that component is approximately Gaussian and nearly independent of  $n_{ch}$ .

In Fig. 5 (left panel) we show the result of subtracting reference  $S_0(y_t)$  from the  $y_t$  spectra in Fig. 1 (right panel) divided by  $n_s/n_{ch}$ . We obtain the difference distributions denoted by  $n_h/n_s \cdot H(\hat{n}_{ch}, y_t)$  (the data points connected with dashed curves). Those data represent *all*  $n_{ch}$  dependence of the  $y_t$  spectra relative to fixed reference  $S_0$ . The error bars denote statistical errors, applicable also to the data in Fig. 1. The two vertical dotted lines enclose region  $B$  on  $y_t$  previously defined.  $H(\hat{n}_{ch}, y_t)$  has unit integral by

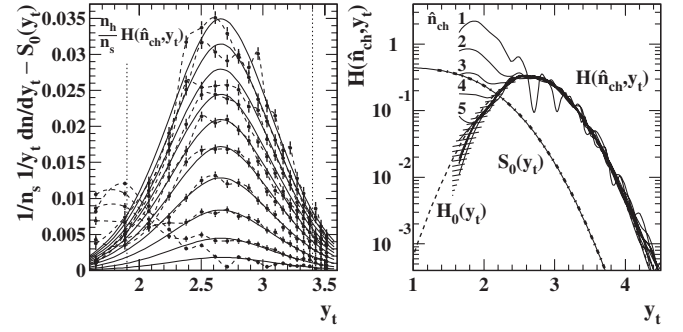


FIG. 5. Left panel: Distributions on  $y_t$  in Fig. 1 (right panel) divided by  $n_s/n_{ch}$  minus reference  $S_0(y_t)$ . Dashed curves in the left panel and solid curves in the right panel are spline fits to guide the eye. The vertical dotted lines enclose region  $B$  previously defined. Right panel: Distributions  $H(n_{ch}, y_t)$  obtained by dividing the curves in the left panel by  $n_h/n_s$ . The dashed curve represents hard reference  $H_0(y_t)$ . The dash-dot curve represents soft reference  $S_0(y_t)$ . The solid curve underlying the dash-dot curve is an error function [19]. The hatched region estimates the systematic error from subtraction of  $S_0$ .

definition, consistent with  $n_s + n_h = n_{ch}$ . The shapes of the data curves are well approximated by the unit-integral Gaussian reference

$$H_0(y_t; \bar{y}_t, \sigma_{y_t}) = A_h(\bar{y}_t, \sigma_{y_t}) \cdot \exp\left[-\frac{1}{2} \left[ \frac{y_t - \bar{y}_t}{\sigma_{y_t}} \right]^2\right], \quad (3)$$

with  $A_h = 0.335 \pm 0.005$ ,  $\bar{y}_t = 2.66 \pm 0.02$ , and  $\sigma_{y_t} = 0.445 \pm 0.005$ . The solid curves represent  $n_h/n_s \cdot H_0$ , with best-fit amplitudes  $n_h(\hat{n}_{ch})/n_s(\hat{n}_{ch})$  plotted in Fig. 7 (right panel, solid dots).  $n_h$  is the multiplicity of the new spectrum component. The data are generally well described by the model, except for the excursions at smaller  $y_t$  for the smaller  $\hat{n}_{ch}$  values.

Dividing the data in Fig. 5 (left panel) by the corresponding best-fit Gaussian amplitudes  $n_h/n_s$  reveals the normalized data distributions  $H(\hat{n}_{ch}, y_t)$  in the right panel. Reference  $S_0(y_t)$ , shown as the dash-dot curve in the right panel, is approximately an error function [19]. The hatched region estimates the systematic error from the  $S_0$  subtraction. Deviations from the  $H_0$  model function (dashed curve) in that panel represent *all* the residual  $n_{ch}$  dependence of the  $y_t$  spectra, i.e., all deviations from the two-component model in Eq. (4) below. Those deviations are plotted in Fig. 6 (left panel) and discussed further in the next section.

The QCD-based power-law trend  $p_t^{-n}$  expected for hard parton scattering would appear in this plotting format as a straight line with negative slope equal to the exponent or “power”  $-n$  [5], since  $y_t \sim \ln(2p_t/m_0)$  at large  $p_t$  makes the plot effectively a log-log plot. Out to  $y_t = 4.5$  or  $p_t = 6$  GeV/c we observe no linear tangential departure from Gaussian model  $H_0$  (dashed parabola) in data  $H(\hat{n}_{ch}, y_t)$ .



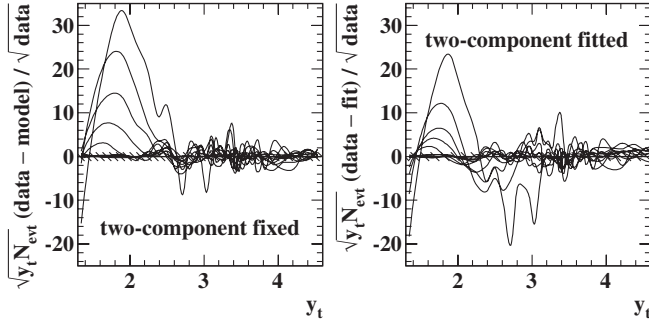


FIG. 6. Left: Relative residuals between data  $y_t$  spectra in Fig. 1 (right panel) and the two-component fixed parametrization, for all multiplicity classes. Right: Relative residuals between data  $y_t$  spectra and two-component free  $\chi^2$  fits, also for all multiplicity classes. The hatched bands represent the expected statistical errors for STAR data.

## VI. TWO-COMPONENT MODEL

The two-component model [10,20] states that the minimum-bias frequency distribution on event multiplicity from relativistic  $p$ - $p$  collisions can be resolved into two components, each approximated by a negative-binomial distribution (NBD) with its own mean and  $k$  parameter. The two components correspond to events with (hard) and without (soft) significant hard parton scatters. That concept can be extended to the possibility that the inclusive  $p_t$  spectrum shape for hard events is different from that for soft events [21]—that the former contains an additional *spectrum* component which we designate the hard component, the complement being then the soft spectrum component. In that interpretation spectra from different multiplicity classes should contain different admixtures of the two spectrum components, and the multiplicity dependence of the spectrum shape may therefore provide a means to isolate those components.

In this section we examine the two-component model in detail. We consider the factorization structure of the model function that has emerged from data analysis, we examine the residuals structure compared to statistical errors, and then test the *necessity* of the fixed-parameter model function by fitting the data with all model parameters freely varying. We finally relate all multiplicities in the model and show that they form a consistent system.

### A. Two-component model function

We have analyzed the multiplicity dependence of  $y_t$  spectra from  $p$ - $p$  collisions without an *a priori* model and have observed a strong  $n_{ch}$  dependence whose functional forms we now summarize. The two-component model of  $y_t$  spectrum structure can be generally represented by the first line of

$$\begin{aligned} 1/y_t dn/dy_t &= s(\hat{n}_{ch}, y_t) + h(\hat{n}_{ch}, y_t) \\ &= n_s(\hat{n}_{ch})S_0(y_t) + n_h(\hat{n}_{ch})H_0(y_t) + \dots, \end{aligned} \quad (4)$$

with unspecified soft and hard spectrum components  $s(\hat{n}_{ch}, y_t)$  and  $h(\hat{n}_{ch}, y_t)$ . What we have inferred from the  $n_{ch}$  dependence of the measured  $y_t$  spectra is the second line, which represents a factorization hypothesis with spectrum components modeled by unit-normal functions  $S_0(y_t)$  and  $H_0(y_t)$  independent of  $n_{ch}$ , ratio  $n_h(\hat{n}_{ch})/n_s(\hat{n}_{ch}) = \alpha\hat{n}_{ch}$ , and constraint  $n_s + n_h = n_{ch}$ . We suggest that the algebraic model in the second line corresponds to the two-component physical model described above and represented by the first line. In the rest of this section we consider the quality and details of the parametrized model in Eq. (4) and test its uniqueness by performing a free  $\chi^2$  fit of the unconstrained model functions to the data.

In the power-law context there is no *a priori* hypothesis for  $n_{ch}$  dependence: each of the ten multiplicity classes in this analysis can be fitted independently with the three-parameter model to produce 30 fit parameters. The corresponding residuals are shown in Fig. 2 (left panel). For the two-component model we could in principle have six free parameters for each  $\hat{n}_{ch}$ , producing 60 fit parameters. However, the algebraic model of Eq. (4) (second line) contains constraints motivated by the requirement of model simplicity which greatly reduce the number of independent parameters. (1) The shapes of unit-integral functions  $S_0(y_t)$  and  $H_0(y_t)$  are independent of multiplicity: each function is determined by only two parameters fixed for all  $n_{ch}$ . (2) The relative normalization of the two components is nearly linearly proportional to the observed multiplicity, as defined by fifth parameter  $\alpha$ . Thus, only five parameters represent all the data in that model. As with the power-law model we compare data to the model on the basis of *relative* fit residuals on  $y_t$ , which provide a more differential and direct assessment of fit quality than the  $\chi^2$  statistic.

### B. Five-parameter fixed model

The residuals in Fig. 6 (left panel) correspond to the function in the second line of Eq. (4) with five optimized parameters held fixed for all  $\hat{n}_{ch}$ . Above  $y_t = 2.7$  the residuals are consistent with statistical fluctuations except for a few sharp structures with amplitude several times the statistical error. Those structures arise from the comparatively low statistics of the Monte Carlo simulations used for background corrections. The Monte Carlo statistical fluctuations appear in these residuals as small-wavelength systematic deviations.

The prominent residuals in  $y_t < 2.7$  for  $\hat{n}_{ch} = 1-4$  (a “third component”) could represent nontrivial  $n_{ch}$  dependence of the soft or hard component or some additional physical mechanism. The end point values at  $y_t = 4.5$  in Fig. 4 (left panel) vary linearly with  $n_{ch}$  to a few percent (open symbols in Fig. 7—right panel), despite the substantial nonlinear excursions at small  $y_t$  of the distributions in Figs. 5 and 6. That *apparent* contradiction suggests that the prominent residuals may represent a change of the hard component at small  $n_{ch}$  which preserves the linear trend of



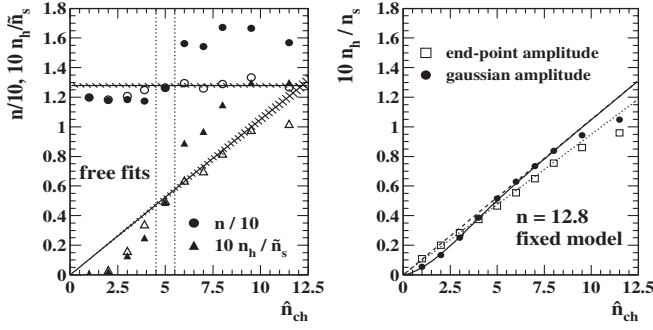


FIG. 7. Left: Parameters  $n$  and  $n_h/\tilde{n}_s$  vs  $\hat{n}_{ch}$  from free  $\chi^2$  fits (solid symbols) in Table I. The open symbols represent similar free fits but with  $\bar{y}_t$  held fixed at 2.65 (see text for discussion). The bands represent the corresponding two-component fixed parametrizations with their stated errors also given in Table I. Right: Ratio  $n_h(\hat{n}_{ch})/n_s(\hat{n}_{ch})$  derived from integrals in Fig. 4 (left panel) (open squares) and from Gaussian amplitudes in Fig. 5 (left panel) (solid dots). The dashed and dotted lines have slopes 0.105 and 0.095, respectively. The solid curve is described in the text.

the integrals. These two-component residuals from the five-parameter fixed model are otherwise *much smaller* than the systematic deviations of the power-law model in Fig. 2 (left panel) with its 30-parameter  $\chi^2$  fit, especially in the large- $y_t$  region where the power-law model should be most applicable.

### C. Two-component free $\chi^2$ fits

To determine whether the algebraic model of Eq. (4) is *necessary* (required by the data), not simply an accident of data manipulation, spectra for  $\hat{n}_{ch} \in [1, 11.5]$  were fitted with the six-parameter function in Eq. (4) using  $\chi^2$  minimization. Spectra  $1/y_t dn/dy_t$  were first normalized by multiplicity estimator  $\tilde{n}_s(\hat{n}_{ch})$  from the fixed parametrization. The coefficients of  $S_0$  and  $H_0$  in the fitting function are then  $n_s/\tilde{n}_s$  and  $n_h/\tilde{n}_s$ . The six parameters ( $n_s$ ,  $\beta_0$ ,  $n$ ,  $n_h$ ,  $\bar{y}_t$ ,  $\sigma_{y_t}$ ) were freely varied for each  $\hat{n}_{ch}$ .

The residuals from the free fits are shown in Fig. 6 (right panel). The fit residuals are comparable to the corresponding fixed-model residuals in Fig. 6 (left panel), even though the free fits include six independent parameters for each of ten  $n_{ch}$  classes for a total of 60 parameters, compared to the fixed model with only five parameters to describe all ten  $n_{ch}$  classes. The residuals for the smaller  $\hat{n}_{ch}$  values show that the free fit attempts to minimize the small- $y_t$  structure (“third component”) in the left panel at the expense of increased intermediate- $y_t$  residuals. The effect on the fit parameters is however modest, as illustrated in Table I.

Table I compares the fixed-model parameter values (fixed) to the results of the six-parameter free fits (fitted) for ten  $\hat{n}_{ch}$  classes. If the hard-component Gaussian on  $y_t$  were not necessary we would expect the  $\chi^2$  fit to converge to the soft-component Lévy distribution as a proxy for the power-law function. The results in Table I indicate that

TABLE I. Two-component  $\chi^2$  fit parameters. The line labeled “fixed” contains the two-component fixed-model parameters. Each fit has  $\nu = 28$  degrees of freedom. The error row applies only to the fixed parametrization. The fit errors are generally smaller than those errors for the last five free-fit rows. Significant systematic effects are discussed in the text.

Fitted	Soft component			Hard component			$\chi^2/\nu$
	$\hat{n}_{ch}$	$n_s/\tilde{n}_s$	$T_0$ (GeV)	$n$	$n_h/\tilde{n}_s$	$\bar{y}_t$ $\sigma_{y_t}$	
1	0.995	0.145	11.97	0.000	—	—	73.3
2	1.001	0.145	11.78	0.002	2.75	0.500	40.4
3	1.001	0.145	11.83	0.013	2.75	0.421	15.0
4	0.996	0.145	11.74	0.025	2.75	0.400	7.36
5	0.994	0.145	12.60	0.049	2.65	0.427	3.14
6	1.001	0.144	15.63	0.089	2.57	0.450	1.09
7	0.999	0.144	15.42	0.097	2.57	0.451	0.62
8	1.005	0.144	16.73	0.115	2.56	0.454	1.18
9.5	1.011	0.143	16.66	0.130	2.56	0.456	0.52
11.5	0.995	0.145	15.69	0.128	2.58	0.460	1.20
Fixed	1.000	0.1445	12.8	$0.0105\hat{n}_{ch}$	2.66	0.445	
Error	0.005	0.001	0.15	$0.0005\hat{n}_{ch}$	0.02	0.005	

most of the free-fit  $S_0$  and  $H_0$  shape parameters remain nearly constant within errors across the full  $\hat{n}_{ch}$  interval. The hard-component Gaussian amplitudes are definitely nonzero and monotonically increasing, consistent with the trends in Fig. 5 (left panel) obtained by subtracting  $S_0(y_t)$  from the normalized spectra in Fig. 1 (right panel).

Figure 7 (left panel) shows trends for the two fit parameters  $n$  and  $n_h/\tilde{n}_s$  which best illustrate the trade-off between soft/power-law and hard components of the model and the necessity of the two-component model. Best-fit values are presented for all  $\hat{n}_{ch}$  classes as the solid symbols (open symbols are discussed below). There are significant systematic deviations of exponent  $n$  from the fixed-model value (hatched band) which are however qualitatively different from the trends in Fig. 2 (right panel). The hard-component amplitude  $n_h/\tilde{n}_s$  also deviates from the linear fixed-model trend, but the trend of monotonic increase is even stronger. The hard component appears to be *more* favored by the free fit than by the fixed parametrization. We discuss the systematic differences between fixed model and free fit in the following paragraphs. However, this fitting exercise does demonstrate that for almost all  $\hat{n}_{ch}$  a two-component model is indeed necessary to describe RHIC  $p$ - $p$  data.

The systematic deviations between free fits and fixed model in Fig. 7 (left panel) are easily understood. We separately consider  $\hat{n}_{ch} < 5$  and  $\hat{n}_{ch} > 5$  (separated by the dotted lines). Generally, there is a strong positive correlation between soft-component exponent  $n$  and hard-component relative amplitude  $n_h/\tilde{n}_s$  originating from the requirement to describe the large- $y_t$  yield. If  $n$  decreases the Lévy distribution tail rises and the amplitude of the hard-component Gaussian amplitude must decrease as well to compensate at large  $y_t$ , and conversely. The systematic

deviations relative to the fixed model for  $\hat{n}_{\text{ch}} < 5$  respond to the presence of the ‘third component,’ which is not a part of the two-component model. To compensate for the additional component in the data the hard-component amplitude is suppressed and  $n$  is reduced by about 10% to provide additional yield from  $S_0$  at small  $y_t$ . The consequence is negative residuals near  $y_t = 2.6$  in Fig. 6 (right panel).

For  $\hat{n}_{\text{ch}} > 5$  a different issue arises. In Fig. 5 (left panel) we have noted previously that the hard-component data peaks are skewed (fall off more rapidly on the low- $y_t$  side) whereas the hard-component model function is a symmetric Gaussian. The difference is most apparent in the running integrals of Fig. 4 (right panel): the dashed model curve lies above the data near  $y_t \sim 2$ . In Fig. 5 (right panel) the hatched region illustrates the region of maximum influence of the  $S_0$  subtraction on the hard component. Because the hard-component data peaks are asymmetric the  $S_0$  subtraction at larger  $y_t$  must be reduced by increasing exponent  $n$  (the small- $y_t$   $S_0$  contribution must remain constant to describe the spectra there). This requires a compensating increase in the hard-component amplitude to fit the larger- $y_t$  part of the spectra, and the Gaussian model function must *shift down* on  $y_t$  (by  $\sim 0.1$  or 5 sigma) and the width increase slightly (0.01 or 2 sigma) to accommodate the *apparent* increased symmetry of the data hard component.

To test that description the free fits were redone with the Gaussian centroid fixed at  $\bar{y}_t = 2.65$ . The open symbols in Fig. 7 (left panel) show the result. The best-fit parameters are now within the error bands of the fixed model, with only modest increase in  $\chi^2/\nu$  (1.69, 1.07, 1.43, 0.95, 1.18, respectively, for  $\hat{n}_{\text{ch}} = 6, \dots, 11.5$  compared to the corresponding values in Table I). The fit residuals in Fig. 6 (right panel) appear identical for the two cases. We emphasize that the mode (most probable point) of the data hard-component peak is near  $y_t = 2.65$ . The downward shift of the *model* peak in the free fit is a consequence of the skewness in the data hard component not described by the fixed model but consistent with measured fragmentation functions from reconstructed jets.

#### D. Two-component multiplicities

In Sec. IVA we adopted a normalization strategy which brought all spectra into coincidence in region A of Fig. 3 (left panel) by defining multiplicity  $n_s \propto \hat{n}_{\text{ch}}$  except for a small deviation linear in  $\hat{n}_{\text{ch}}$ . We then defined reference function  $S_0$  as a limiting case of the spectrum  $n_{\text{ch}}$  dependence and isolated a second component  $H_0$  by subtracting the fixed reference from all spectra. The amplitude of  $H_0$  relative to the reference is defined by ratio  $n_h/n_s \propto \hat{n}_{\text{ch}}$ . The representation to that point is (physics) model independent, derived only from the observed spectrum  $\hat{n}_{\text{ch}}$  dependence: the reference is  $\propto \hat{n}_{\text{ch}}$  and the second compo-

nent is  $\propto \hat{n}_{\text{ch}}^2$ . That difference is the *underlying basis* for distinguishing the two components.

In this section we have identified the two algebraic spectrum components with the components of a physical model of soft and hard parton scattering and subsequent fragmentation to detected particles. We distinguish four event multiplicities: (1) the observed multiplicity  $\hat{n}_{\text{ch}}$  or uncorrected number of particles with  $p_t > 0.2$  GeV/c in the STAR angular acceptance which serves as an event-class index, (2) the corrected and  $p_t$ -extrapolated multiplicity  $n_{\text{ch}}$ , (3) the ‘‘soft-component’’ multiplicity  $n_s$  and (4) the ‘‘hard component’’ multiplicity  $n_h$ , with  $n_s + n_h = n_{\text{ch}}$ . We now examine the self-consistency of the multiplicities in our two-component model in the context of real spectrum properties, including efficiencies and acceptances.

Soft multiplicity  $n_s$  is estimated by function  $\tilde{n}_s(\hat{n}_{\text{ch}}) = [2.0 \pm 0.02(\text{rel}) \pm 0.2(\text{abs})]\hat{n}_{\text{ch}}[1 - (0.013 \pm 0.0005)\hat{n}_{\text{ch}}]$ . The 1% error applies to the relative or spectrum-to-spectrum normalization relevant to this differential analysis, whereas the 10% error applies to the common normalization of all spectra. As noted, coefficient 0.013 is determined by requiring that corrected spectra normalized by  $\tilde{n}_s$  approximately coincide within region A of Fig. 3 (left panel) for all  $\hat{n}_{\text{ch}}$ . The factor 2 is determined by requiring that after correction, extrapolation with  $S_0$  to  $p_t = 0$  and normalization with  $n_{\text{ch}}$  all spectra in Fig. 1 integrate to unity. In the first column of Table I deviations of  $n_s/\tilde{n}_s$  from unity are consistent with the 1% error estimate.

The hard fraction  $n_h/n_s = \alpha\hat{n}_{\text{ch}}$  is estimated by two methods. In the first method we determine the Gaussian amplitudes required to fit the data distributions in Fig. 5 (left panel). Those amplitudes give the solid Gaussian curves compared to data in that plot and are plotted as the solid points in Fig. 7 (right panel). The linear trend (dashed line) corresponds to slope  $\alpha = 0.0105 \pm 0.0005$ . The solid curve passing precisely through the points is  $n_h(\hat{n}_{\text{ch}})/n_s(\hat{n}_{\text{ch}}) = \{(0.0105\hat{n}_{\text{ch}})^{-10} + (0.005\hat{n}_{\text{ch}}^{1.5})^{-10}\}^{-1/10}$ , the errors on the coefficients being  $\pm 0.0005$ . The nonlinearity of that curve is related to the non-Gaussian small- $y_t$  structure for small values of  $\hat{n}_{\text{ch}}$  (third component).

In the second method we note that the distributions in the left panel of Fig. 4 are running integrals of data distributions in Fig. 5 (left panel). The amplitudes of those integrals at end point  $y_t = 4.5$ , plotted as open squares in Fig. 7 (right panel), also estimate ratio  $n_h/n_s$ . They vary nearly linearly (dotted line) with slope  $\alpha = 0.0095 \pm 0.0005$ . Reduction of  $\alpha$  from 0.0105 for the Gaussian amplitudes to 0.0095 for the integral endpoints results from small deviations of the data peaks from the  $H_0$  Gaussian model at small  $y_t$  evident in Fig. 5. The data are slightly skewed in a manner consistent with measured fragmentation functions. The model Gaussians are matched to the data at and above the data peak mode or

most probable point. The integral of any data peak is therefore expected to be slightly less than that of the corresponding model function. Both methods suggest saturation of the hard-component amplitude at larger  $\hat{n}_{\text{ch}}$ .

Consistency of the soft and hard multiplicity estimators within the two-component model can be established by the following argument: Tracking inefficiencies produce the same fractional changes for all  $\hat{n}_{\text{ch}}$  and are represented by factors  $\epsilon_s$  and  $\epsilon_h$  for soft- and hard-component yields. The corrected spectra are extrapolated to  $p_t = 0$  with soft model  $S_0$ . The fraction of  $S_0$  falling above  $p_t = 0.2$  GeV/c (within the  $p_t$  acceptance) is represented by  $\gamma$ . The hard component identified in this analysis falls entirely within the  $p_t$  acceptance. The *observed* multiplicity is then given by  $\hat{n}_{\text{ch}} \equiv \gamma \epsilon_s n_s + \epsilon_h n_h$ , whereas the corrected and extrapolated spectra integrate to true multiplicity  $n_{\text{ch}} = n_s + n_h$ . The expression for  $\hat{n}_{\text{ch}}$  above can be rearranged to solve for  $n_s$  in the first line below,

$$\begin{aligned} n_s &\simeq \frac{\hat{n}_{\text{ch}}}{\gamma \epsilon_s} \left\{ 1 - \frac{\epsilon_h}{\gamma \epsilon_s} \cdot \alpha \hat{n}_{\text{ch}} \right\} \quad \text{predicted,} \\ \tilde{n}_s &= 2\hat{n}_{\text{ch}} \{ 1 - 0.013\hat{n}_{\text{ch}} \} \quad \text{observed,} \end{aligned} \quad (5)$$

whereas the second line is the estimator inferred from the data. By integrating reference  $S_0$  we determine that  $\gamma = 0.7$ : 70% of the reference spectrum is within the acceptance  $p_t > 0.2$  GeV/c. Tracking efficiencies  $\epsilon_s$  and  $\epsilon_h$  are both approximately 70%, and we have determined from the data (running integrals) that  $\alpha \sim 0.0095$ . We therefore have  $1/\gamma \epsilon_s \sim 2$  and  $\epsilon_h/\gamma \epsilon_s \cdot \alpha \sim 0.0135$ , establishing the consistency (predicted  $\leftrightarrow$  observed) of the two-component multiplicities. Coefficient 0.013 is identified as  $\alpha/\gamma$ , and the trend of  $n_s$  is defined by ratio  $n_h/n_s = \alpha \hat{n}_{\text{ch}}$ . We thus close the circle, demonstrating quantitatively how increase with  $\hat{n}_{\text{ch}}$  of the hard-component contribution to the spectrum forces  $n_s$  to decrease relative to  $\hat{n}_{\text{ch}}$  in compensation, why  $\tilde{n}_s$  contains the negative term, and what its magnitude must be.

## VII. $\langle p_t \rangle$ SYSTEMATICS

Another aspect of the two-component model is the variation of  $\langle p_t \rangle$  (inclusive mean  $p_t$ ) with  $\hat{n}_{\text{ch}}$ . Estimation of  $\langle p_t \rangle$  for spectra with incomplete  $p_t$  acceptance requires either a model fit or direct integration of data with extrapolation. The power-law function for  $p_t$  distributions  $1/p_t dn/dp_t = A/(1 + p_t/p_0)^n$ , with  $\langle p_t \rangle = 2p_0/(n - 3)$ , has been used previously to extract  $\langle p_t \rangle$  values from corrected  $p_t$  spectra [8].  $\langle p_t \rangle$  can also be determined by direct integration of the experimental  $p_t$  spectra, with extrapolation to  $p_t = 0$  by a suitable model function. Finally, the two-component fixed-model function obtained in this analysis can provide a parametrization of  $\langle p_t \rangle(n_{\text{ch}})$ .

The running multiplicity integral  $n(\hat{n}_{\text{ch}}, y_t)$  is defined by Eq. (1), with the data extrapolated over  $p_t[0, 0.2]$  GeV/c by reference  $n_s(n_{\text{ch}})S_0(p_t)$ . Running integral  $p_t(\hat{n}_{\text{ch}}, y_t)$  can

also be defined for transverse momentum  $p_t$  by including an extra factor  $p_t(y_t)$  in the integrand of Eq. (1). The ratio  $\langle p_t \rangle(\hat{n}_{\text{ch}}, y_t) = p_t(\hat{n}_{\text{ch}}, y_t)/n(\hat{n}_{\text{ch}}, y_t)$  is then a function of  $y_t$  for each value of  $\hat{n}_{\text{ch}}$ , and  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  is the limit of that function as  $y_t \rightarrow \infty$ .  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  is thus determined by direct integration of  $p_t$  or  $y_t$  spectra.

A changing mixture of soft and hard components may cause  $\langle p_t \rangle$  to vary with  $n_{\text{ch}}$ . The  $\langle p_t \rangle$  values for individual components are obtained by direct integration of model functions  $S_0$  and  $H_0$ :  $\langle p_t \rangle_{\text{soft}} = 0.385 \pm 0.02$  GeV/c and  $\langle p_t \rangle_{\text{hard}} = 1.18 \pm 0.01$  GeV/c. A two-component *analytic expression* for  $\langle p_t \rangle$  is then given by

$$\langle p_t \rangle(\hat{n}_{\text{ch}}) = \left\{ 0.385 \frac{n_s(\hat{n}_{\text{ch}})}{n_{\text{ch}}} + 1.18 \frac{n_h(\hat{n}_{\text{ch}})}{n_{\text{ch}}} \right\} \text{ GeV/c,} \quad (6)$$

with  $n_h/n_s = \alpha \hat{n}_{\text{ch}}$  and  $n_s + n_h = n_{\text{ch}}$ .

In Fig. 8  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  values inferred from power-law fits to corrected STAR spectra are represented by open circles, consistent with a 200 GeV UA1 power-law analysis plotted as solid triangles [8], but inconsistent at smaller  $\hat{n}_{\text{ch}}$  with the two-component result from this analysis plotted as solid points. The two-component data were obtained with the  $n_h/n_s$  values plotted as solid dots in Fig. 7 (right panel). The solid line represents the two-component analytic expression for  $\langle p_t \rangle$  in Eq. (6) with  $\alpha = 0.0095$ . The UA1 results for 900 GeV [8] are plotted as open triangles. The dotted line corresponds to Eq. (6) with  $\alpha = 0.015$ .  $\langle p_t \rangle$  values obtained by *direct integration* of the extrapolated spectra are represented by the open squares. The hatched region represents the common uncertainty in all means due to uncertainty in the particle yield in  $p_t < 0.2$  GeV/c.

The  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  values in Fig. 8 obtained by direct integration of extrapolated spectra provide the best estimate of the physical trend. The results at 200 GeV for direct

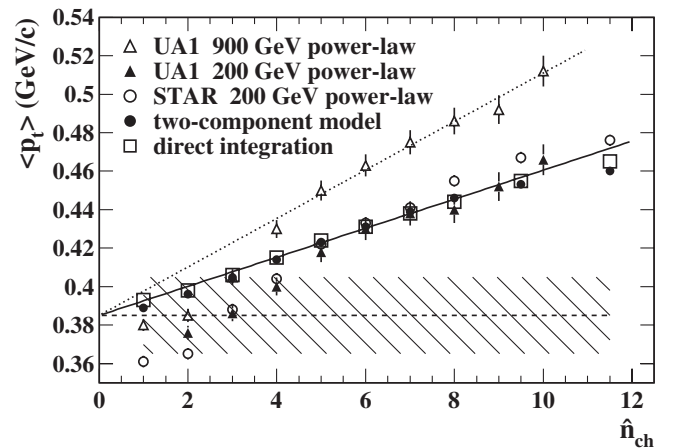


FIG. 8.  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  derived from the two-component  $H_0$  Gaussian amplitudes (solid dots), from the running integrals (open squares), and from power-law fits to STAR and UA1 data (open circles, triangles). The solid and dotted lines correspond to Eq. (6) with  $\alpha = 0.0095$  and  $0.015$ , respectively.



integration, the two-component model and power-law fits are consistent within errors for  $\hat{n}_{\text{ch}} > 4$ . The notable deviation of the power-law results from the two-component linear trend for  $\hat{n}_{\text{ch}} < 5$  can be explained by the three-component structures at small  $y_t$  and small  $\hat{n}_{\text{ch}}$  in Fig. 5 (left panel). Those structures strongly influence (bias) extrapolation of the power-law function into the unmeasured region in  $p_t < 0.2$  GeV/c so as to overestimate the inferred yield there (nominally 30% of the total spectrum). The overestimate at small  $p_t$  produces a sharp reduction of  $\langle p_t \rangle(\hat{n}_{\text{ch}})$  values inferred from power-law fits. The additional yield at small  $y_t$  in Fig. 5 itself corresponds to  $\langle p_t \rangle \sim 0.4$  GeV/c, and thus cannot physically lower the composite  $\langle p_t \rangle$  below  $\langle p_t \rangle_{\text{soft}} = 0.385$  GeV/c. These  $\langle p_t \rangle$  results demonstrate that the UA1 data are sensitive to the small- $y_t$  and small- $\hat{n}_{\text{ch}}$  structures revealed in this analysis when the more integral spectrum measure  $\langle p_t \rangle$  is used.

### VIII. ERRORS

The statistical errors for the basic  $y_t$  spectra in Fig. 1 are best indicated by the error bars on the difference distributions of Fig. 5 (left panel). That figure also compares the pointwise statistical errors to the hard-component structure inferred in this study, which is statistically well determined for all  $n_{\text{ch}}$  classes. Monte Carlo calculations of background corrections with full detector response simulation are computer intensive. Because of limited statistics the statistical fluctuations in the Monte Carlo data used for background corrections are injected into the corrected data spectra as visible systematic errors: long-wavelength systematic error is reduced at the expense of increased short-wavelength random “systematic” error. Those errors are apparent as the nonstatistical short-wavelength structures in Figs. 2 and 6. The systematic uncertainties in the corrected spectra can be divided into  $n_{\text{ch}}$ -dependent and  $n_{\text{ch}}$ -independent uncertainties.

$n_{\text{ch}}$ -independent systematic uncertainties include uncertainties in the corrections for tracking efficiency, backgrounds (mainly weak decays), and momentum resolution. Systematic spectrum corrections for this analysis were 20% or less, except for the lowest two  $p_t$  bins where they increased to 40%. Statistical errors for the systematic corrections were typically less than 1% (except as noted above for the background corrections). We estimate the uncertainties in the systematic corrections as 10% of the correction values. The total uncertainty for the systematic corrections is then less than 2% above  $p_t = 0.4$  GeV/c. The UA1 corrected  $n_{\text{ch}}$ -inclusive  $p_t$  spectrum for 200 GeV  $\bar{p}$ - $p$  collisions [8] agrees with the corresponding inclusive spectrum from the present analysis at the 2% level.

$n_{\text{ch}}$ -dependent systematic errors could result from  $n_{\text{ch}}$ -dependent tracking inefficiencies. However, track detection and  $p_t$  measurement in this analysis required no reference to other tracks or a fitted event vertex, thus minimizing any  $n_{\text{ch}}$  dependencies. In effect, each track

was treated in isolation independent of its relationship to any event, except for the timing requirement with the CTB. The tracking efficiencies for low-multiplicity (1–4) and high-multiplicity ( $> 4$ ) events integrated over the  $p_t$  acceptance were found to be consistent to 3%, with a 1% statistical error. We take that as an estimate of the  $n_{\text{ch}}$ -dependent systematic uncertainty.

The main source of systematic uncertainty in the shape of the hard-component structures isolated in Fig. 5 is the definition of  $S_0$  as the lowest element of the regular sequence in Fig. 4 (left panel).  $S_0$  is a rapidly decreasing function in the interval  $y_t = 1.6$ –3. The main effect of varying either  $\beta_0$  or  $n$  in  $S_0$  is to change the magnitude of  $S_0$  in that interval, shape changes being secondary. It is consistent within the two-component context to require that (1) component  $H(n_{\text{ch}}, y_t)$  be non-negative, placing an upper bound on  $S_0$  in Fig. 5 and (2) that any  $n_{\text{ch}}$ -independent aspect of the distributions in Fig. 5 be minimized, determining a lower bound. Those criteria place stringent constraints on  $S_0$  already in  $y_t \sim 1.6$ –2, limiting systematic offsets at  $y_t = 2$  to  $\pm 0.002$ , the allowed range rapidly decreasing above that point according to the  $S_0$  curve in Fig. 10 (right panel). The systematic uncertainty estimate corresponding to those trends is represented by the hatched region in Fig. 5 (right panel).

The nonstatistical power-law fit residuals in Fig. 2 are as much as 30 times the statistical error. One of the findings of this study is that the power-law model function is inappropriate for these  $p_t$  spectra. Systematic uncertainties for the fit parameters are therefore not meaningful.

The fitting uncertainties for the fixed-model parameters are given at the bottom of Table I. Those uncertainties are meaningful relative to the fitting procedure defined in the two-component model context. The ability of that model to describe the data is apparent in Fig. 6 (left panel). The only significant residuals correspond to a low- $y_t$  spectrum element (for  $\hat{n}_{\text{ch}} = 1$ –4) deliberately omitted from the two-component model. One source of systematic uncertainty in those parameters is whether the fixed-model prescription forces a certain result by excluding some other which may better describe the data.

To test that possibility a free  $\chi^2$  fit with all model parameters varying was conducted. The difference in the two cases is summarized in Table I and Fig. 7 (left panel). In particular, there are substantial differences in the Lévy exponent and the hard-component amplitude for the free fit depending on whether the position of the hard-component Gaussian is constrained or not. When the Gaussian position is constrained the free fit and the fixed model agree within the systematic uncertainties in the latter. The differences in the unconstrained fit are traced to significant departures of the shape of the hard-component data peak from the symmetric Gaussian peak shape: the model function could be further refined by adding a skewness (expected for fragmentation functions) to improve the stability of the fits.



However, it is not our purpose to develop a complex representation of  $y_t$  spectra, but rather to demonstrate the essential two-component aspects of the spectra with the simplest possible model function. The differences in fit parameters in Fig. 7 (left panel) can therefore be taken as a generous estimate of the systematic uncertainty in the fixed-component parametrization.

### IX. IDENTIFIED PARTICLES

Model functions  $S_0$  and  $H_0$  derived from this analysis of unidentified particles represent physical spectrum components  $S$  and  $H$  for several hadron types, mainly  $\pi$ ,  $K$ , and  $p$ . Two questions emerge: (1) to what extent do  $S_0$  and  $H_0$  correspond to individual hadron types, and (2) to what extent does the  $n_{\text{ch}}$  dependence of the  $p_t$  spectrum truly separate two physical components  $S$  and  $H$ ? The soft component of one hadron species may have significant  $n_{\text{ch}}$  dependence which could be misinterpreted as the hard component of another species, or of the combination of unidentified hadrons in this study.

We can obtain some answers to those questions from  $n_{\text{ch}}$ -inclusive spectrum studies of identified hadrons.  $p_t$  spectra for  $\sqrt{s_{NN}} = 200$  GeV  $p$ - $p$  collisions have been measured for identified pions, kaons, and protons [22]. Because  $\hat{n}_{\text{ch}} \sim 1$  and the  $p_t$  acceptance was  $[0.3, 3]$  GeV/c for that analysis the measured multiplicity-inclusive  $p_t$  spectra are reasonably described by Lévy distribution  $S_0$ , especially the kaon and proton spectra. The common Lévy exponent for the three species is  $n = 16.8 \pm 0.05$ , compared to  $n = 12.8 \pm 0.15$  measured in this analysis for unidentified hadrons. The slope parameter for identified pions is  $T = 0.145 \pm .001$  GeV, whereas for both kaons and protons  $T = 0.23 \pm 0.005$  GeV, compared to  $T = 0.1445 \pm 0.001$  GeV for unidentified hadrons in this analysis.

The trend of  $S_0$  with hadron species is easily understood. Addition of the “hotter”  $K$  and  $p$  spectra to the “cooler” pion spectrum flattens the unidentified-hadron composite at larger  $p_t$ , reducing the exponent of  $S_0$  to  $n = 12.8$ . At smaller  $p_t$  the pion fraction dominates the composite spectrum, and the unidentified-hadron slope parameter is the same as the pion slope parameter. The effect of the heavier hadrons on the composite spectrum is mainly to reduce the Lévy exponent from the larger physical value common to all three hadron species.

Information on the  $n_{\text{ch}}$  dependence of  $p_t$  spectra for identified particles is limited. A preliminary analysis of  $K_S^0$  and  $\Lambda$   $p_t$  spectra up to 4 GeV/c [23] suggests that the  $n_{\text{ch}}$  dependence of both spectra can be described by a modest (5%) reduction of  $n$  with increasing  $n_{\text{ch}}$ . That trend can be compared to the free  $\chi^2$  fit results for  $S_0$  in Table I as shown in Fig. 7 (left panel):  $n$  increases by about 25% over the measured  $\hat{n}_{\text{ch}}$  range. That increase is traced to an attempt by the model to accommodate a skewness of the

hard component in the data, not a true variation in the soft component.

### X. PYTHIA MONTE CARLO

A similar analysis of  $p$ - $p$  collisions from the Pythia Monte Carlo [24] reveals substantial deviations from data. We studied default Pythia-V6.222 and Pythia “tune A” (increased initial-state radiation and multiple soft parton scatters relative to the default) with parameters derived from studies of the underlying event in triggered jet events [25]. In Fig. 9 (left panels) we show Pythia  $p_t$  spectra normalized to unit integral and soft reference  $S_0$  (dash-dot curves) determined by the same criteria applied to STAR data. Those plots can be compared to Fig. 3 (left panel). In Fig. 9 (right panels) we show the results of subtracting reference  $S_0$  from the normalized spectra in the left panel divided by  $n_s/n_{\text{ch}}$ . Those plots can be compared to Fig. 5 (left panel). The dashed curves are the hard component  $H_0$  for STAR data divided by 10 to provide a reference.

The  $S_0$  parameters for Pythia-V6.222 in the upper panels are  $T_0 = 0.147$  GeV and  $n = 23$ . The large value of  $n$

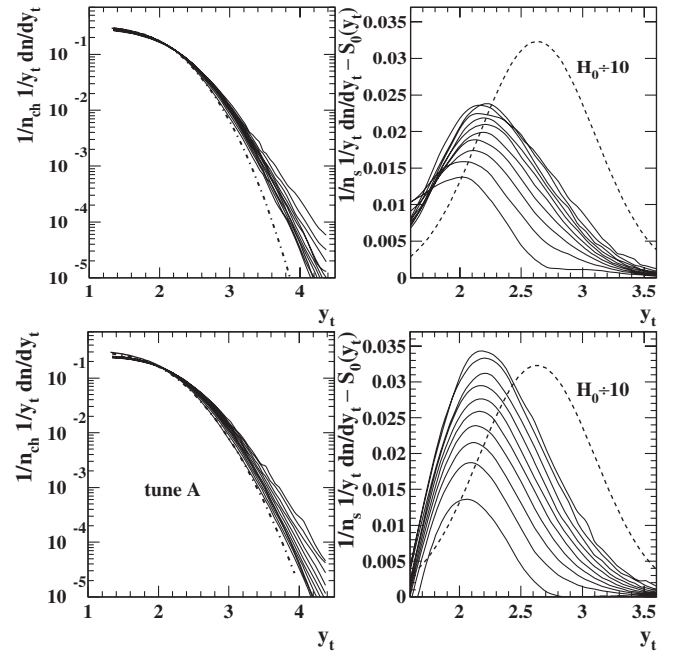


FIG. 9. Two-component analysis applied to Pythia Monte Carlo data with the same multiplicity classes as for STAR data. The left and right panels may be compared with Figs. 3 (left panel) and 5 (left panel), respectively. Dashed curves in the right panels represent the STAR data hard component for  $\hat{n}_{\text{ch}} \sim 11$  ( $H_0/10$ ). Dash-dot curves in the left panels represent soft component  $S_0$  optimized for each Monte Carlo configuration: Pythia V6.222 default with parameters  $T_0 = 0.147$  GeV and  $n = 23$  (upper panel); Pythia tune A with parameters  $T_0 = 0.137$  GeV and  $n = 14$  (lower panel).

implies that the Pythia soft component is nearly Maxwell-Boltzmann, in sharp contrast to RHIC data. The exponent is strictly limited to a large value by the Pythia data for  $y_t < 2.5$ . The  $S_0$  parameters for Pythia tune A in the lower panels are  $T_0 = 0.137$  GeV and  $n = 14$ . The smaller value of  $n$  is comparable to the value 12.8 observed for RHIC data.

The hard-component yield for Pythia is generally a factor of 2 to 3 less than the data (most apparent above  $y_t = 2.7$ ), broader and peaked at a smaller value of  $\bar{y}_t$ . Pythia-V6.222 shows saturating of the hard-component amplitude with increasing  $\hat{n}_{ch}$ , whereas tune A shows a more uniform and significantly greater rate of increase. The large Gaussian-shaped offset *common to all curves* and centered at  $y_t \sim 2$  is also not observed in the data. That structure cannot be accommodated by the Lévy distribution. The two Pythia Monte Carlos thus exhibit some features which agree qualitatively with experimental data but are quantitatively different. Tune A is closer to data than the default for soft and hard components, but the  $n_{ch}$ -independent Gaussian offset near  $y_t = 2$  persists and is not observed in the data.

## XI. DISCUSSION

A description of  $p$ - $p$  collisions in terms of soft and hard components is natural at RHIC energies where significant hard parton scattering occurs but the underlying event [25] is still relatively simple. The two-component model of nuclear collisions can be applied to (1) the event-frequency distribution on  $n_{ch}$  (two or more negative-binomial distributions) [10,20], (2) the dependence of  $\langle p_t \rangle$  on  $n_{ch}$  [8,16,26], (3) triggered jet correlations on  $(\eta, \phi)$  (correlations from soft and hard event classes) [21], and (4) the  $n_{ch}$  dependence of the  $p_t$  or  $y_t$  spectrum shape [16]. The common theme is the relation of hard parton scattering to event multiplicity in the context of a “soft” underlying event. This paper emphasizes analysis type (4)—the study of the  $n_{ch}$  dependence of the spectrum shape on transverse momentum  $p_t$  and transverse rapidity  $y_t$ .

Model functions  $S_0(y_t)$  and  $H_0(y_t)$  in Eq. (4) can be viewed as the lowest-order elements of a perturbative expansion of the spectrum shape. Multiplicities  $n_s(\hat{n}_{ch})$  and  $n_h(\hat{n}_{ch})$  can be interpreted as estimating the mean numbers of soft- and hard-component particles per-event for a given  $\hat{n}_{ch}$ . The claim of simplicity for the two-component fixed model is supported by the small number of parameters, the simplicity of the model functions, the demonstration of necessity in Sec. VIC, and the demonstration with residuals plots that there is no additional information in the spectra (aside from the small- $y_t$  “third component” which may represent additional physics).

We cannot rule out additional components or changes in the shapes of *physical* components  $S$  and  $H$ . Each should be  $n_{ch}$ -dependent at some level, but the present analysis

indicates that within the observed  $\hat{n}_{ch}$  interval any such dependence is near the level of statistical error. A change in  $S$  is suggested by the  $n$  dependence of the free  $\chi^2$  fits in Fig. 6 (left panel). However, that behavior may simply be due to a coupling of soft and hard amplitudes in the free fit, with no physical significance.

A significant change in  $H$  is expected at larger  $\hat{n}_{ch}$  based on known jet physics: larger fragment multiplicities are produced by more energetic partons, with fragment distributions shifted to larger  $y_t$  [13]. Thus, the mean and width of  $H$  should increase with  $\hat{n}_{ch}$  at some point, but such changes are not observed beyond statistics within the  $y_t$  and  $\hat{n}_{ch}$  acceptances of this study. Apparently, the multiplicity increase in this analysis is dominated by increased *frequency of events with a single hard scattering* within a multiplicity class rather than bias toward more energetic partons. That scenario is consistent with the two-component model of [10].

The soft-component Lévy distribution  $S_0 \equiv A_s/(1 + \beta_0(m_t - m_0)/n)^n$  [17] is similar in form to power-law function  $A/(1 + p_t/p_0)^n$ . However, the physical interpretations are quite different. The Lévy distribution describes a nominally exponential function with a control parameter (e.g., slope parameter) which undergoes Gaussian-random fluctuations. Inverse exponent  $1/n$  then measures the *relative variance*  $\sigma_\beta^2/\beta_0^2$  of the control parameter [27]. In the limit  $1/n \rightarrow 0$  the Lévy distribution on  $m_t$  becomes a true Maxwell-Boltzmann distribution. Those properties suggested the Lévy distribution as a reference function for this analysis. Ironically, the “power-law” function in the form of a Lévy distribution describes the soft-component, not hard parton scattering. The Lévy parameters can be interpreted in the context of an ensemble of hadron emitters with random transverse speeds, thermal radiation from moving sources as described by the Cooper-Frye formalism [28]. The expected QCD hard-scattering power-law trend is not evident in the data out to  $p_t \sim 6$  GeV/c.

In Fig. 2 (right panel) we plot exponent  $n$  values from power-law fits to data (solid points) and to the two-component fixed model (open circles) for the full range of  $\hat{n}_{ch}$ . The latter procedure simulates a power-law fit to data with no small- $y_t$  excursions or third component and illustrates the effect of those features on the exponent. The range of variation of the power-law exponent, in contrast to the two-component fixed model, and the substantial effect of the third component further illustrate that the power-law parametrization is sensitive to aspects of spectra inconsistent with its theoretical motivation, making fit results difficult to interpret physically.

The Gaussian shape of  $H_0(y_t)$  inferred from this analysis can be compared with *fragmentation functions* from jet analysis of  $p$ - $p$ ,  $e$ - $p$ , and  $e$ - $e$  collisions plotted on logarithmic variable  $\xi_p \equiv \ln\{p_{jet}/p_{fragment}\}$ , which also have an approximately Gaussian shape [29] explained in a QCD context as the interplay of parton splitting or branching at

larger  $p_t$  and the nonperturbative cutoff of the branching process at smaller  $p_t$  due to gluon coherence [30,31]. The Gaussian parameters are predicted by the pQCD modified leading-log approximation (MLLA) [32]. The hard component obtained in this analysis then represents not fragmentation functions from reconstructed large- $E_t$  jets but rather the average of a *minimum-bias* ensemble of fragmentation functions dominated by *low- $Q^2$*  parton scatters ( $Q < 10$  GeV). In that context  $H_0$  represents *minimum-bias* partons dominated by *minijets* [33]. A previous study of small- $E_t$  clusters in 200 GeV  $p$ - $p$  collisions [34] suggested that semihard parton scattering or gluon radiation from projectile constituent quarks could produce substantial small- $p_t$  structure in hadron spectra similar to the hard component of this study.

A recent analysis of  $p_t$  spectra in the interval 0.3–10 GeV/c for identified particles in  $p$ - $p$  and  $d$ -Au collisions [35] used the relativistic-rise particle identification scheme to extend the spectra with very good statistics to large  $p_t$ . That paper compared the spectra to several NLO pQCD calculations and compared the  $m_t$  spectra of pions and protons. It concluded that there is a transition region from soft to hard particle-production processes at  $p_t \sim 2$  GeV/c in inclusive particle production, which would appear to contradict the present results. However, the identified-particle spectra in that study below  $p_t = 2.5$  GeV/c are from a previous study [22] in which the point-to-point systematic errors and the statistical errors are quite large, the latter due to the small acceptance of the prototype ToF detector. The ToF-based studies of multiplicity-averaged  $p$ - $p$  collisions are therefore not sensitive to the hard-component structure reported in this paper, the great majority of which falls below 2.5 GeV/c. The present study takes a new approach by comparing large-statistics inclusive-hadron spectra in several multiplicity bins. Since the hard component is relatively enhanced in high-multiplicity events we are able to extend our investigation of the hard component to low  $p_t$  by studying the trend of that enhancement.

The relative frequency of hard scatters in  $p$ - $p$  collisions is described by the fifth model parameter  $\alpha \sim 0.01$ , representing the nearly linear dependence of  $n_h/n_s$  on  $\hat{n}_{ch}$ . We relate the hard-component amplitude to the frequency of hard collisions ( $f \equiv$  number of hard collisions per NSD  $p$ - $p$  collision) as  $n_h(\hat{n}_{ch}) = \alpha \hat{n}_{ch} n_s(\hat{n}_{ch}) = f(\hat{n}_{ch}) \cdot \bar{n}_{mj}$ , with mean true event multiplicity  $\bar{n}_{ch} = 2.5$  in one unit of pseudorapidity and mean minijet multiplicity  $\bar{n}_{mj} = 2.5 \pm 1$  [36]. We then estimate the observed frequency of hard scatters in  $\sqrt{s} = 200$  GeV  $p$ - $p$  collisions as  $f = \bar{n}_h/\bar{n}_{mj} = 0.012 \pm 0.004$  observed hard scatters per NSD  $p$ - $p$  collision per unit of pseudorapidity. In that interpretation multiplicity  $\hat{n}_{ch}$  serves as a “trigger” for hard parton scattering, determining the fraction of hard-scattering events in a given multiplicity class and thus the relative amplitude of the hard spectrum component.

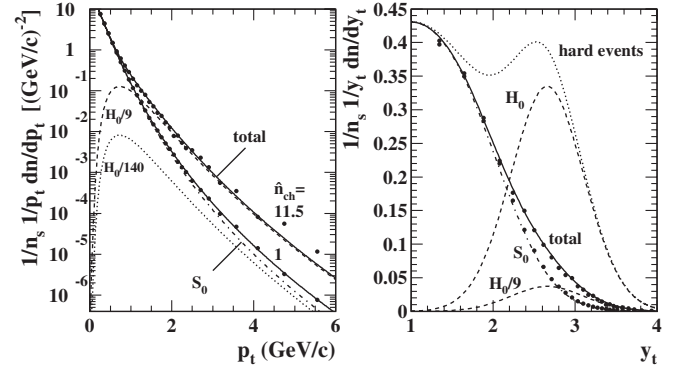


FIG. 10. Decomposition of inclusive  $p_t$  spectra into a soft component represented by a Lévy distribution on  $m_t$  and a hard component represented by a Gaussian on  $y_t$ . Dashed curves  $H_0/9$  correspond to data for  $\hat{n}_{ch} = 11.5$ , while the dotted curve  $H_0/140$  corresponds to data for  $\hat{n}_{ch} = 1$  (cf. Fig. 1). The dash-dot curves are soft reference  $S_0$ , and the solid curves are the totals of soft and hard components for the model. The dotted curve in the right panel estimates the shape of the inclusive  $y_t$  distribution for those  $p$ - $p$  collisions containing at least one minimum-bias hard parton scatter (hard events).

Model functions  $S_0$  and  $H_0$  on  $p_t$  and  $y_t$  are summarized in Fig. 10, which can be compared with Figs. 1, 3, and 5.  $H_0/9 \Rightarrow n_h/n_s = 0.11$  is compared to data for  $\hat{n}_{ch} = 11.5$  and illustrates the role of the hard component in the measured spectra with sufficient amplitude to be visible in a linear plotting format (right panel). Similarly,  $H_0/140 \Rightarrow n_h/n_s = 0.007$  is compared to data for  $\hat{n}_{ch} = 1$ . Those coefficients are consistent with the measured hard-component Gaussian amplitudes for  $\hat{n}_{ch} = 1$  and 11.5 (cf. Fig. 7—right panel).

Collisions in the event ensemble containing at least one semihard parton scatter within the detector acceptance should have similar yields of soft and hard components (assuming an average minijet multiplicity of 2.5). The average  $y_t$  spectrum for such hard events is illustrated by  $S_0 + H_0$ , shown as the dotted curve in Fig. 10 (right panel). We cannot isolate such hard events in an unbiased manner, but we can infer their structure by extrapolating the  $n_{ch}$  trends determined in this analysis.

The left panel of Fig. 10 indicates the loss of visual sensitivity to spectrum structure when spectra are plotted on  $p_t$ . The hard component can appear to be a continuation of the soft component, whereas in the right panel the two components are clearly separate functional forms.  $y_t$  provides a more balanced presentation of structure resulting from hard-scattered parton fragmentation, yet does not compromise study of the soft component, which is well described by a simple error function on  $y_t$  [19]. The transverse and longitudinal fragmentation systems undergo similar physical processes and should therefore be compared in equivalent plotting frameworks. Just as  $y_z$  is preferred to  $p_z$  we prefer  $y_t$  to  $p_t$ .

## XII. SUMMARY

In conclusion, we have studied the event multiplicity  $n_{\text{ch}}$  dependence of high-statistics transverse momentum  $p_t$  or transverse rapidity  $y_t$  spectra from  $p$ - $p$  collisions at  $\sqrt{s} = 200$  GeV. We have determined that the power-law model function fails to describe the spectra for any  $n_{\text{ch}}$ , exhibiting large nonstatistical deviations from data. An earlier UA1 study reporting satisfactory power-law fits to data seems contradictory. However, it is statistically consistent with the present study because the UA1 data were derived from a much smaller event sample. We have analyzed the shapes of the spectra with a running-integral technique and determined that the spectra can be described precisely by a simple five-parameter model function. The algebraic model can in turn be related to a two-component physical model of nuclear collisions.

The power-law function motivated by pQCD expectations for hard parton scattering better describes the soft component in the form of a Lévy distribution on  $m_t$  (two parameters). We observe for the first time that the hard component is well described by a Gaussian distribution on transverse rapidity  $y_t$ , with shape approximately independent of multiplicity (two parameters). The hard-component multiplicity *fraction* increases almost linearly with event multiplicity (the fifth model parameter). A detailed comparison of (data – model) residuals from the two-component fixed model and from free fits with all two-component model parameters varied confirms that the two-component fixed model is required by the data.

The hard component may represent fragments from transversely scattered partons. The shape is consistent with fragmentation functions observed in LEP and PETRA  $e^+e^-$  and FNAL  $p$ - $\bar{p}$  collisions. The stability of the hard-component shape with event multiplicity suggests that a Gaussian distribution on  $y_t$  is a good representation of minimum-bias parton fragments. The relative abundance of soft and hard components at any  $y_t$  of course depends on  $y_t$  and  $n_{\text{ch}}$ , but most of the hard-component yield falls below 2.5 GeV/c. There is evidence for a small but significant third component at smaller  $y_t$  and smaller  $n_{\text{ch}}$ . Comparison with the Pythia Monte Carlo reveals qualitative differences from data.

## ACKNOWLEDGMENTS

We thank the RHIC Operations Group and RCF at BNL, and the NERSC Center at LBNL for their support. This work was supported in part by the Offices of NP and HEP within the U.S. DOE Office of Science; the U.S. NSF; the BMBF of Germany; IN2P3, RA, RPL, and EMN of France; EPSRC of the United Kingdom; FAPESP of Brazil; the Russian Ministry of Science and Technology; the Ministry of Education and the NNSFC of China; IRP and GA of the Czech Republic, FOM of the Netherlands, DAE, DST, and CSIR of the Government of India; Swiss NSF; the Polish State Committee for Scientific Research; SRDA of Slovakia, and the Korea Sci. and Eng. Foundation.

## APPENDIX: SYMBOL DEFINITIONS

Below is a list of symbols and their definitions as used in this paper.

- $y_t$ : transverse rapidity, replaces transverse momentum  $p_t$  to provide improved visual access to fragment distributions
- $\hat{n}_{\text{ch}}$ : observed event multiplicity in the detector acceptance, also the event-class index
- $n'_{\text{ch}}$ : efficiency- and acceptance-corrected multiplicity in the detector acceptance
- $n_{\text{ch}}$ : corrected and  $p_t$ -extrapolated or “true” multiplicity in the detector angular acceptance
- $n_s$ : soft-component multiplicity in the acceptance
- $\tilde{n}_s$ : particular function of  $\hat{n}_{\text{ch}}$  used to estimate  $n_s$
- $n_h$ : hard-component multiplicity in the acceptance:  $n_s + n_h = n_{\text{ch}}$
- $\alpha$ : hard-component coefficient:  $n_h/n_s \sim \alpha \hat{n}_{\text{ch}}$
- $S_0$ : unit-normal functional form of the soft component (Lévy distribution on  $m_t$ )
- $N_0$ : running integral of soft reference  $S_0$
- $H_0$ : unit-normal functional form of the hard component (Gaussian distribution on  $y_t$ )
- $A, p_0, n$ : power-law model parameters
- $A_s, \beta_0, n$ : soft-component Lévy distribution parameters,  $1/\beta_0 = T_0$ , the slope parameter
- $A_h, \bar{y}_t, \sigma_{y_t}$ : hard-component Gaussian parameters

- 
- [1] J. Adams *et al.* (STAR Collaboration), Phys. Rev. Lett. **91**, 172302 (2003).
  - [2] S. S. Adler *et al.* (PHENIX Collaboration), Phys. Rev. Lett. **91**, 072301 (2003); B. B. Back *et al.* (PHOBOS Collaboration), Phys. Lett. B **578**, 297 (2004); I. Arsene *et al.* (BRAHMS Collaboration), Phys. Rev. Lett. **91**, 072305 (2003).
  - [3] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Phys. Rep. **97**, 31 (1983).
  - [4] X. N. Wang and M. Gyulassy, Phys. Rev. D **44**, 3501 (1991).
  - [5] M. J. Tannenbaum, Eur. Phys. J. C **43**, 329 (2005); J. Phys.: Conf. Ser. **27**, 1 (2005).
  - [6] X. N. Wang, Eur. Phys. J. C **43**, 223 (2005).



- [7] R. C. Hwa and C. B. Yang, Phys. Rev. C **67**, 034902 (2003); V. Greco, C. M. Ko, and P. Lévai, Phys. Rev. Lett. **90**, 202302 (2003); R. J. Fries, B. Müller, C. Nonaka, and S. A. Bass, Phys. Rev. Lett. **90**, 202303 (2003).
- [8] C. Albajar *et al.* (UA1 Collaboration), Nucl. Phys. **B335**, 261 (1990).
- [9] R. Blankenbecler, S. J. Brodsky, and J. F. Gunion, Phys. Lett. B **42**, 461 (1972); R. F. Cahalan, K. A. Geer, J. Kogut, and L. Susskind, Phys. Rev. D **11**, 1199 (1975).
- [10] A. Giovannini and R. Ugoccioni, Phys. Rev. D **59**, 094020 (1999).
- [11] Transverse rapidity  $y_t = \ln\{(m_t + p_t)/m_0\}$  in a longitudinally comoving frame ( $p_z = 0$ ) is equivalent in form to longitudinal rapidity  $y_z = \ln\{(E + p_z)/m_0\}$  in a transversely comoving frame ( $p_t = 0$ ).
- [12] D. V. Minh and P. Carruthers, Phys. Rev. Lett. **31**, 133 (1973).
- [13] M. Z. Akrawy *et al.* (OPAL Collaboration), Phys. Lett. B **247**, 617 (1990); D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **68**, 012003 (2003). O. Biebel, P. Nason, and B. R. Webber, hep-ph/0109282.
- [14] K. H. Ackermann *et al.* (STAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 624 (2003).
- [15] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics* (Van Nostrand, Princeton, NJ, 1962), 3rd ed.
- [16] T. Alexopoulos *et al.* (E735 Collaboration), Phys. Lett. B **435**, 453 (1998).
- [17] G. Wilk and Z. Włodarczyk, Physica A (Amsterdam) **305**, 227 (2002).
- [18] Soft-component  $n = 12.8 \pm 0.15$  from this study compares with power-law  $n = 11.8 \pm 0.4$  from  $p$ - $\bar{p}$  collisions [8] where the power-law exponent represents soft *and* hard components: flatter  $p_t$  distributions require smaller  $n$  values.
- [19] The Lévy distribution on  $y_t$  is well approximated by an error function:  $S_0(y_t; y_{t0}, \xi_{y_t}) = A(y_{t0}, \xi_{y_t}) \times \{1 + \text{erf}[-(y_t - y_{t0})/\xi_{y_t}]\}/2$ . For these data  $A = 0.454 \pm 0.005$ ,  $y_{t0} = 2.03 \pm 0.02$ , and  $\xi_{y_t} = 0.76 \pm 0.03$ .
- [20] S. G. Matinyan and W. D. Walker, Phys. Rev. D **59**, 034022 (1999).
- [21] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **65**, 072005 (2002).
- [22] J. Adams *et al.* (STAR Collaboration), Phys. Lett. B **616**, 8 (2005).
- [23] M. Heinz (STAR Collaboration), J. Phys. G **31**, S1011 (2005).
- [24] T. Sjöstrand, Comput. Phys. Commun. **82**, 74 (1994); T. Sjöstrand, L. Lönnblad, S. Mrenna, and P. Skands, hep-ph/0308153.
- [25] T. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **65**, 092002 (2002); D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **70**, 072002 (2004); R. Field and R. C. Group (CDF Collaboration), hep-ph/0510198.
- [26] D. Kharzeev and M. Nardi, Phys. Lett. B **507**, 121 (2001).
- [27] J. Adams *et al.* (STAR Collaboration), nucl-ex/0408012.
- [28] F. Cooper and G. Frye, Phys. Rev. D **10**, 186 (1974).
- [29] M. Z. Akrawy *et al.* (OPAL Collaboration), Phys. Lett. B **247**, 617 (1990).
- [30] C. P. Fong and B. R. Webber, Nucl. Phys. **B355**, 54 (1991).
- [31] Y. L. Dokshitzer, V. S. Fadin, and V. A. Khoze, Phys. Lett. B **115**, 242 (1982).
- [32] C. P. Fong and B. R. Webber, Phys. Lett. B **229**, 289 (1989); Y. L. Dokshitzer, V. A. Khoze, C. P. Fong, and B. R. Webber, Phys. Lett. B **273**, 319 (1991).
- [33] K. Kajantie, P. V. Landshoff, and J. Lindfors, Phys. Rev. Lett. **59**, 2527 (1987); A. H. Mueller, Nucl. Phys. **B572**, 227 (2000); G. C. Nayak, A. Dumitru, L. D. McLerran, and W. Greiner, Nucl. Phys. **A687**, 457 (2001).
- [34] C. Albajar *et al.* (UA1 Collaboration), Nucl. Phys. **B309**, 405 (1988).
- [35] J. Adams *et al.* (STAR Collaboration), Phys. Lett. B **637**, 161 (2006).
- [36]  $\bar{n}_{mj} \approx 2.5$  at  $\sqrt{s} \approx 200$  GeV is estimated from jet multiplicity systematics in T. Affolder *et al.* (CDF Collaboration), Phys. Rev. Lett. **87**, 211804 (2001); see also the survey of fragment multiplicities from  $e^+ - e^-$  collisions in I. M. Dremin and J. W. Gary, Phys. Rep. **349**, 301 (2001).